

Health research and market failure

The British government needs to respond urgently to the charge that attempts to introduce an 'internal market' into the operation of the health service are undermining the country's clinical research base.

FEW will have failed to see the irony that the Clinton administration's efforts at health-care reform, designed to increase the social spread of those able to afford adequate medical treatment, have coincided with equally controversial moves by the British government to put its own health system, long heralded as a model of universal health-care, on a more market-oriented basis. In both countries, the biomedical research community has been legitimately concerned that its own needs have not — at least initially — figured high on the priority list of political reformers. And in both cases, it is becoming clear that special precautions are needed if such research is to be protected from reformers.

Costly

In Britain, these dangers have been highlighted by a report published last week by the United Kingdom Coordinating Committee on Cancer Research (UKCCCR), a body that represents the main organizations funding cancer research (see page 499). The report draws attention to a specific threat to long-term clinical trials arising from the British government's moves to put the allocation of health resources on a 'purchaser/provider' basis. The government wants to make those responsible for delivering health services (whether hospital doctors or general practitioners) directly responsible for ensuring that this is achieved in the most cost-effective way. The UKCCCR rightly points out that the more such 'purchasers' are concerned with the immediate effectiveness of the services they are paying for — meaning in practice by the amount of medical care delivered to their patients — the less resources and attention they are likely to spare for long-term, communal goals needing to be addressed through research.

Two immediate responses can be made to the UKCCCR's charges. One is that public money spent on supporting the clinical trials of new drugs could be considered as a subsidy to the pharmaceutical companies that will eventually market (and profit from) such drugs, both nationally and internationally. The UKCCCR and its supporters argue that the unique research opportunities for broad-based clinical trials offered by the National Health Service (NHS) have been partly responsible for the strength of Britain's pharmaceutical companies. It could be argued that the scale of such subsidies is no longer appropriate, and that public funds would be better spent on the basic biomedical research that will produce the drugs of the future.

The second response is that there is an element of the

charges that is inevitably self-serving. The changes in the NHS have, not surprisingly, generated widespread resentment among those who find they are being less generously funded from the public purse than in the past. But the government's concern to ensure that public money is effectively spent is a legitimate one. And often the most effective way of doing this is to put responsibility for it squarely on those who stand to lose most (and are also likely to complain the loudest) if this goal is not achieved. Much as researchers may resent a 'payment by results' regime, it is more satisfactory that one that makes an open-ended commitment to good ideas, however inefficiently pursued.

Yet neither counter-argument is sufficient to deny the validity of the UKCCCR's main complaint, namely that the main thrust of the government's reforms — that an internal market for health services is sufficient to provide the most cost-effective health system — is already proving severely inadequate to address the pressing funding needs of many high-value research projects. In the wider market economy, basic research is generally accepted, even by Britain's conservative government, as a 'common good' whose support is a collective responsibility. The same argument must be applied to the 'internal markets', such as those in health, which the government is so keen to encourage.

Limitations

The government's first move should be to publish as soon as possible the report it has commissioned from Anthony Culyer of the University of York, however embarrassing its criticisms of current philosophy, to demonstrate its willingness to address the problem head-on. Its second should be to come forward with proposals that demonstrate an awareness of the limitations of a market-led approach in supporting a healthy science base in the NHS, and the imagination to implement the necessary steps to improve the situation. Moves along the lines of the UKCCCR's suggested 'top-slicing' of health service funds (intriguingly similar to recommendations for 'top-slicing' medical insurance contributions in the United States, as proposed by Senators Tom Harkin and Mark Hatfield and others) could be one solution. Others exist. But they each require acknowledgement of the strengths of collective action, precisely the spirit that the government is seeking to replace with its doctrine of individual responsibility. Bridging the gap is a top priority; restoring the flagging morale of health service scientists comes close behind. □



Time to resist earmarks

A US Congressman's fight against earmarked funds needs and deserves more public support from the universities.

CONGRESSMAN George Brown (Democrat, California) continues to pursue his single-handed battle against academic earmarking with tenacity and some imagination. Despite recent setbacks, he has just opened up a new front in his campaign, demanding that energy secretary Hazel O'Leary resist the diversion of money from the Human Genome Project into a senator's pet project.

Brown's letter to O'Leary serves the public well, highlighting the earmarking process at its most insidious. It asks her to consider the particular case of \$1.1 million cut from her department's contribution to the Human Genome Project to pay for new Apple Power PCs at the Oregon Health Science University.

The Oregon school does no research at all for the energy department — yet according to Brown it has received \$96 million in earmarks, mostly from the energy pot, since 1983. Senator Mark Hatfield (Republican, Oregon) is the ranking republican member of the Energy and Water Development appropriations subcommittee.

The energy department spends plenty of time agonizing over what to cut, but makes no attempt whatsoever to evaluate the scientific value of the earmarks themselves. Brown has told O'Leary that she can spend money on earmarks if she wants, "but I expect you to have a more compelling reason than a desire to placate a distinguished senior Senator on the Appropriations Committee".

The earmarkers, being a surreptitious bunch, rarely make their case in public, but they basically have three arguments in their favour.

The worst one is that they send money to small schools, while a proper peer review process favours the likes of Harvard and Massachusetts Institute of Technology. This is like the Mafia defending extortion on the grounds that it boosts the economy. If the weaker schools are treated unfairly by peer review — and a recent study by the Government Accounting Office found that they are not — then the solution rests with better peer review procedures, not with the arbitrary favours of the earmarking system.

Slightly more convincing is the appropriators' charge that the federal budget allocates next to nothing for new university buildings, and that only earmarks can get them built. Last month's Science in the Public Interest document from the White House should have stated how this perennial problem is going to be resolved. Instead it promised that the National Science and Technology Council would "develop options". It must do so with greater urgency than has so far been evident.

The third and most telling argument for earmarking is simply that a cynical public expects nothing less and nothing more. When the powerful and allegedly corrupt former chairman of the House Ways and Means committee, Dan Rostenkowski, faced a severe primary challenge in his

Chicago district earlier this year, he duly arranged for the federal government to deliver new helicopter ambulances to local hospitals during the campaign — which he won hand-somely. George Brown has actually been criticised in California for complaining about earmarking instead of getting on with the business of abusing his position in Congress to bring home the bacon. The public, some would argue, gets the government it deserves.

The setback for Brown's anti-earmarking campaign came last month (see *Nature* 369, 694; 1994) when John Murtha (Democrat, Pennsylvania), chairman of the House defense appropriations subcommittee and earmarker par excellence, threatened to cut the defence department's \$1.8 billion university research budget in half. If this was intended to scare the universities into silence on the earmarking issue, it must not succeed.

Congressman Brown, of course, has his own agenda: his attacks on appropriators are designed, in part, to nurture the influence of the Science, Space and Technology authorization committee which he chairs. Nonetheless he has occupied the high moral ground. It is time for the scientific community to join him there: otherwise, it will be correctly judged as just another special interest group with its nose in the public trough. □

Germany and the bomb

The likelihood that a cartel is forming to sell bomb-grade plutonium exported to Germany is a cause for alarm.

LAST week German officials seized some 500 grams of radioactive plutonium 239 from luggage aboard a flight from Moscow to Munich. It was the third and largest recent seizure of plutonium that appears to have originated in Russia (despite the denials of Russian officials) and suggests that an organized plutonium cartel is at work.

So far, officials have acknowledged two other seizures of weapons-grade material. Plutonium 239 was found in a garage in Germany in May; enriched uranium 235 was seized in Bavaria in June. German officials claim that scientific analysis shows that the fissionable material comes from Russia, while Russian spokesmen say there have been no losses of plutonium from any of their facilities, with the possible exception of missing medical-grade material. But the radioactive isotopes used in medicine would not, in any case, be useful for building bombs and bombs are what are at issue here.

The end of the Cold War has lead to the faulty presumption that the world is now a safer place. Witness the apparent decline in weapons-building activity by the US Department of Defense. But in fact, the possibility that terrorists or unstable small governments may acquire the ability to build nuclear warheads is as destabiling a threat as there could be. And it should go to the top of the list of concerns among all of the nations of the world. This is not just the making of a good novel of international intrigue. It threatens the stability (and viability) of the world. □

Patent fights over hepatitis C test kits reverberate around the world

Sydney. Australia has become the latest battleground in a continuing and acrimonious conflict over patents awarded to the US biotechnology company Chiron Corporation of Emeryville, California, covering the rights to the DNA sequence of hepatitis C virus (HCV) and to tissue cultures and vaccines developed from that sequence.

The patent has been controversial since Chiron first applied to patent offices around the world for patent protection for its hepatitis C screening kits in November 1987. Since then Chiron has been awarded patents in various countries, including Australia, Germany, the Netherlands, Japan, the United Kingdom and Japan, although the US patent has not as yet been granted.

But the breadth of the patent remains

sharply contested. Scientists in Australia, for example, where Chiron is suing one of its competitors, International Murex Technologies Corporation, for infringing the patent, claim that the granting of the patent is already hampering research on the virus.

Last week, Stephen Locarnini, head of the Victorian Infectious Diseases Reference Laboratory at Fairfield Hospital in Melbourne, Victoria, said that one sponsor had pulled out of a research project after inspecting the Chiron patent.

Locarnini admits that the resulting shortfall in funds had been made up through another programme. But he adds that the patent granted to Chiron is so broad that it is difficult for small investors to put money into research on the hepatitis C virus.

Australia is not the only country in which legal battles are being fought. Last month, for example, Organon Teknika nv, a division of the Dutch multinational chemical company Akzo Nobel nv, announced an appeal against a court injunction stopping it from distributing hepatitis screening kits until a full hearing has been held of charges of patent infringement brought against it by Chiron.

The order was made by the Hague district court at the demand of Chiron, following last year's award of a patent to Chiron by the European Patent Office. The patent application had been strongly contested by Organon Teknika, which imports HCV test kits made by the US company United Biomedical Inc.

Organon distributes kits in the United Kingdom, the Netherlands and Germany, and is being sued by Chiron in each of these countries. Murex is also being sued by Chiron in the United Kingdom and Italy, as well as Australia, over its anti-HCV test, manufactured by its subsidiary, Murex Diagnostics. Murex says that, like Organon, it is seeking to have Chiron's HCV patents revoked.

The Australian patent, which is similar to that filed in most other countries, contains more than 30 clauses. The first of these covers most of the DNA sequence for the HCV, while the remainder cover techniques for tissue culturing and vaccine development from that sequence.

Larry Kurtz of Chiron claims that, without the reward offered by the patent system, it is doubtful whether companies such as Chiron would make the substantial investments needed to solve the problem posed by the HCV.

The case against Murex for violation of Chiron's patent is expected to come to court in Australia in the middle of next year. Chiron has already won a similar challenge to its patent in the United Kingdom during hearings on suits filed against both Murex and Organon Teknika.

An initial hearing resulted in the patent being ruled as valid, except for clauses on tissue culturing and vaccine development which were subsequently removed from the UK patent. A second hearing resulted in a decision that the defendants had infringed the patent. The final hearing is due to take place in November, when Chiron expects to be awarded damages against the defendants, and an injunction stopping the distribution of their products.

But the defendants still intend to appeal, arguing that the patent is invalid because it is in the public interest to allow more than one company to market HCV tests.

Mark Lawson

Strasbourg goes for top potential

Paris. The dream of school physics classes everywhere, the world's largest Van de Graaff generator (see right), built at a cost of FFr100 million (US\$18.5 million) and planned eventually to generate a potential of 35 million volts (MV), opened last month at the Centre de Recherches Nucléaires in Strasbourg, France.

But although the cost of the so-called Vivitron may have raised some hairs, the main aim of the



machine is to study the structure of atomic nuclei by accelerating ions. Indeed, Strasbourg believes the new machine will reinforce its ambition to become the European centre of excellence in this area.

The Vivitron started up at just 16.6 MV, and this charge will be progressively increased to 35 MV through next year. The final voltage will be bigger than that of the most powerful existing machines of this type, in particular a 26-MV accelerator at Oak Ridge, Tennessee, in the United States, and a 22-MV in the United Kingdom.

By focusing large but controlled energies on nuclei, Vivitron will cause collisions that produce nuclei with unusual compositions of excited internal states. In particular, it will be used to study 'super-deformation', a phenomenon discovered at Daresbury in 1986.

This is an excited and relatively stable state in which rapidly rotating nuclei deform and acquire an elliptical form. Physicists

say that such nuclei act as a 'laboratory' in which laws governing the structure of the nucleus can be explored in a simplified form.

Usually observed in nuclei of mass around 130, 150 or 190, superdeformation is detected by the emission of specific fingerprints of gamma rays. Identifying these against the noisy background is the job of the FFr60 million Anglo-French detector Eurogam II, which is coupled to Vivitron.

Three other detectors — DEMON, a FFr7 million neutron detector, and ICARE and CHARISSA, both charged particle detectors — will later be coupled to Vivitron. The successor to Eurogam II, which is already the world's highest resolution gamma-ray spectrophotometer, is being planned by France, Britain, Germany, Italy and the Scandinavian countries.

Strasbourg is pitching to host the FFr140 million EUROBALL detector. But the Italian city of Legnaro, near Venice, would also like to host the machine.

Declan Butler

AIDS chief promises a shift towards basic research . . .

Yokohama. The United States will reorientate its fight against AIDS to encourage more broad-based basic research, according to the head of the US AIDS research effort, outlining the new policy at the 10th International AIDS conference in Yokohama last week.

William E. Paul, director of the Office of AIDS Research at the US National Institutes of Health (NIH), was appointed to the post six months ago after having worked for several years at NIH as an immunologist (see *Nature* 367, 587; 1994).

Paul said that it is "time to chart a new course for the future, avoiding the twin perils of an unreflective allegiance to the status quo and an impulsive acceptance of easy answers and solutions". He reaffirmed the NIH's continued commitment to clinical research, which receives a very large proportion of the total NIH AIDS research budget.

But he said that this research will be restructured to make it "more complete, coherent and cost-effective". And he called on the pharmaceutical industry to shoulder more of the costs of large phase-three clinical trials.

Echoing views recently expressed by Bernard Fields of Harvard Medical School (see *Nature* 369, 95; 1994), Paul described basic research as "the engine that will drive the entire AIDS research enterprise forward". In particular, he said that the current inadequacy of treatments for HIV infection were due to "wide gaps in our understanding" of the pathogenesis both of HIV disease and of the virus itself.

"A vigorous programme of creative investigator-initiated work is essential", said Paul, pledging to increase the resources available for support of such work. "I will

use all my powers to persuade a broader pool of basic scientists to join us in this effort."

At the same time, Paul warned that the basic research effort should not be overmanaged. Research administrators need to remember that breakthroughs would come from insights that cannot be planned. "Command science is no more likely to succeed than command economics."

NIH's drug development would also be reorientated, he said, to put "special emphasis" on therapies based on new molecular targets that will require greater understanding of the virus.

Paul said that agents that target the virus itself, such as the drug AZT, have been of limited success in treatment of HIV infection. A broader approach would therefore be adopted to develop new agents that intervene in the immunopathogenesis of the disease as well as to find new ways, such as gene therapy, to attack the virus itself.

Paul emphasized his office's commitment to vaccine development, despite the recent setback when the NIH had withheld approval of phase-three trials of two gp120-based candidate vaccines. But large-scale testing would proceed only when reliable scientific evidence suggested a "reasonable degree of promise".

In another reflection of the NIH's increased commitment to basic research, he said that in order to provide insight into vaccine development, support will be increased for studies of immunological resistance to infection with HIV and related lentiviruses. One such example he quoted was the Nairobi Sex Workers Study, which is examining a group of women who, despite repeated exposure to HIV, have resisted infection.

David Swinbanks

. . . as Japanese taunt comes under attack

Yokohama. It is not only Japanese politicians who are capable of making public statements that upset the outside world. Shortly after the opening of the 10th International Conference on AIDS, Yuichi Shiokawa, the head of the conference organizing committee, came under fire for blaming the spread of AIDS in Japan on the "Americanization" of Japanese sexual morality.

Shiokawa had originally made the remark at a press conference held at the Ministry of Health and Welfare last month. At a further press conference after last week's opening ceremony, he was asked to explain what he had meant. Shiokawa said he meant that the "sexual behaviour of the Japanese has become more open and free than the traditional

one-to-one relationship".

Participants at the conference were upset by Shiokawa's remark, not only for its linking of AIDS to sexual morality — or the fact that it is questionable whether Japanese sexual behaviour has ever been traditionally one-to-one — but because the majority of AIDS cases in Japan have been caused by the use of contaminated blood products.

Shiokawa served on a government AIDS committee in 1983 that has recently been under attack for its failure at the time to allow the rapid introduction of heat-treated blood products (see *Nature* 367, 584; 1994). As a result of the delay, thousands of Japanese haemophiliacs were infected with HIV, and they remain the majority of Japan's AIDS cases. □

Work on telescope to continue in Chile as options are viewed

Munich. The council of the European Southern Observatory (ESO) last week agreed to continue construction work on the DM500-million (US\$330 million) Very Large Telescope (VLT), even though a dispute over the site of the telescope on the top of Mount Paranal has not yet been fully resolved (see *Nature* 368, 676; 1994).

At an extraordinary meeting held in Munich, the council agreed that the major mechanical components of the first VLT 8.2-metre unit telescope (there will be four such units in all) will be shipped to Chile later this year, enabling the project to remain on schedule for operation in 2000.

But, as guarantees of the long-term security of ESO's access to the site in northern Chile have still not been assured, the council also decided to continue its investigations of an alternative site in Namibia, in southern Africa, in case negotiations with the Chilean government fail.

Mount Paranal has already been levelled and foundations are being laid in preparation for the construction of the VLT, which will be the world's largest optical telescope. But the plans came under threat two years ago when a local family disputed ownership of the site (see *Nature* 368, 676; 1994).

ESO argues that it was granted immunity from local laws by an agreement with the Chilean government signed in 1963. Chilean officials say they accept ESO's general assessment, but are insisting on a formal amendment to the original agreement to cover the situation at Paranal.

The detailed wording of the proposed amendment is proving difficult to negotiate, but ESO's director, Riccardo Giacconi, says he is "very optimistic" that this can be rapidly sorted out.

Giacconi says that he is encouraged by his successful negotiation of resolutions to other disputes that have dogged ESO's activities in Chile for several years, including greater viewing times offered to Chilean astronomers at all ESO's telescopes in Chile.

But Giacconi says that it is "only pragmatic" to consider other options. ESO is now studying a site in Namibia on the flat top of the Gamsberg mountain, which is at present owned by Germany's Max Planck Society. Gamsberg was originally considered as a possible site for the VLT but Mount Paranal was seen as offering better overall viewing conditions.

Last month ESO started monitoring the viewing conditions in Namibia in case the site is needed as a fallback. But Giacconi hopes that ESO can stay in Chile. "It would be preposterous for science if we were forced to build the VLT at what is not the best possible site," he says.

Alison Abbott

US energy moves signal budget fights ahead

Washington. Attempts by Republican members of the US Congress to place an upper limit on research spending at the Department of Energy (DoE) have given a warning of what the research community can expect if — as is widely predicted — the Republican party makes substantial gains in November's congressional elections.

Representative Robert Walker (Republican, Pennsylvania) has been leading an effort to limit non-defence research and environmental restoration work at the DoE at this year's level of \$4.29 billion until 1998. This move would eliminate any increases for inflation, and force all new spending — such as that proposed for new fusion and high-energy physics facilities — to be taken out of other work covered by the limit.

Walker's bid to limit DoE spending has taken the form of an amendment to the Hydrogen, Fusion and High Energy and Nuclear Physics Authorization Act. The move has angered many Democrats, who point out that overall federal spending is already capped by existing law.

But Walker and his conservative backers believe that such an overall limit still encourages Congress to spend money in favoured areas — such as scientific research — without considering where the money will come from.

The House of Representatives was this week due to debate a version of the bill that already included a cap of \$3 billion on research into energy supply (although some of the bill's supporters assumed this would be removed from the bill in conference with the Senate).

Walker's amendment extends the scope of the cap to cover high-energy and nuclear physics as well; if adopted, the result would be that the proposed expansion in these fields would have to be offset elsewhere.

Sherwood Boehlert (Republican, New York) was due to offer a compromise amendment raising the cap by \$50 million for three years. The extra money would be used to fund the high-energy physics work proposed earlier this year by a panel led by Sidney Drell of the Stanford Linear Accelerator Center (see *Nature* 369, 266; 1994).

The authorization bill does not directly set research budgets — that is done each year by appropriations committees. But observers of the science policy scene in Washington say that the importance of Walker's amendment is as a pointer to the future for research funding.

Research scientists who have spent recent years complaining about budgets that have risen only at close to the rate of inflation may now see budgets frozen in cash terms, and policed by a more fiscally conservative Congress.

The Republicans are unlikely to win control of the House of Representatives in November, when all its members and one-third of senators are up for re-election. But they are likely to return as a larger and more conservative group, easily capable of winning a majority in alliance with conservative Democrats on spending issues.

Some of the resulting pain is likely to be felt by environmental and physical scientists. But it is expected that their biomedical colleagues will to some degree be protected by the stronger bipartisan support for their work.



Walker: bidding to limit DoE spending.

Once the House has dealt with the authorization bill, its fate will rest with Senator Bennett Johnston (Democrat, Louisiana), who steered a bill authorizing fusion alone through the Senate last year and must now decide if he wants to proceed with the

House's wider bill.

Johnston was originally keen on a fusion bill to ensure that the United States had a coherent plan for fusion, and prevent it from 'wasting' \$650 million on the planned Tokamak Physics Experiment (TPX) at Princeton Plasma Physics Laboratory, New Jersey (see *Nature* 367, 669; 1994). He has not so far revealed his position on the House of Representatives' proposals for high-energy physics, which support Drell's call for US participation in Europe's Large Hadron Collider (LHC).

If Johnston is enthusiastic, a bill could be passed before the present Congress holds its final sitting in October. That would be good news for physicists (especially if the spending caps are knocked out), as it would offer some protection to fusion and high-energy physics programmes in what are expected to be difficult years ahead by formally endorsing them in law. It would also lend some purpose to the long, hot August sitting that has been forced upon Congress by the Clinton administration's efforts at health-care reform.

Colin Macilwain

Biosphere 2 wins growing approval

Washington. Biosphere 2, the ecological laboratory built in Arizona by Ed Bass, a Texan billionaire, has replaced its scientific director and is planning a new scientific programme in partnership with Columbia University, New York.

Officials say the changes mark a new beginning for the much-criticized laboratory, which includes the world's largest hermetically sealed greenhouse. Biosphere 2 has asked a number of prominent scientists to produce reports by the end of the year on its research potential in a dozen different scientific disciplines, ranging from ecological modelling to nutrition.

The scientific programme will be overseen by a committee of four members: Michael Crow, vice-provost of Columbia; Wallace Broecker, of Columbia's Lamont-Doherty Earth Observatory; Bruno Marino, Biosphere's new scientific director; and Stephen Bannon, acting chief executive of the private company that runs Biosphere 2 for Bass.

The committee will be given advice by a panel of scientists chaired by Michael McElroy, chairman of the Earth sciences department at Harvard University.

Marino was also previously a member of Harvard's Earth sciences department, and replaces John Corliss, who will continue working on ecological modelling at Biosphere 2. Under an agreement with Columbia, Biosphere 2 will pay an undisclosed sum for its assistance in setting up its new programme.

John Mutter, the interim director of Lamont-Doherty, says that what he describes as an "extraordinary facility" has in effect been released from its previous mission to become a serious scientific facility. "That presents a great opportunity for our scientists," he says.

Asked if Biosphere 2's track record would make it hard to attract support for research, Mutter says that "the only thing we can do is put that behind us. We were not involved in previous missions."

The seven-strong "crew" currently manning Biosphere 2 will emerge earlier than planned next month, cutting a link with the emphasis on human "survival" which critics had attacked as a gimmick, and opening the way for a more conventional programme of scientific experiments at the \$150-million facility (see *Nature* 368, 88; 1994).

Bass sacked the more holistically minded members of Biosphere 2's management earlier this year, and has been soliciting advice from outside scientists on how best to exploit the project's considerable scientific potential (see *Nature* 368, 576; 1994).

Biosphere 2 and Columbia hope to build a scientific programme that will attract substantial funding from outside sources such as the National Science Foundation. Given the physical attributes of the facility, and the Clinton administration's keenness to support environmental research, that should not prove too difficult.

Colin Macilwain

Nuclear critics set sights on French plutonium pledge . . .

Munich. The French nuclear fuel company Cogema has offered to store either nuclear waste or separated plutonium for German utilities as an inducement to them not to pull out from long-term contracts to reprocess spent fuel from nuclear reactors.

But leaked details of the offer have prompted a storm of protest from anti-nuclear activists. These point out that Cogema's suggestion would be illegal under current legislation, which requires the plutonium separated from reprocessed fuel to be returned to its country of origin.

The environmental group Greenpeace has demanded an inquiry into the affair, accusing Cogema of negotiating in defiance of French law.

Reprocessing separates the one per cent of plutonium that exists in spent fuel, and this can be used as a component of mixed-oxide (MOX) fuel, which can be burnt in some power stations (see below).

During the 1980s, German utilities signed contracts with Cogema and its UK counterpart British Nuclear Fuels (BNFL) for reprocessing their nuclear waste in the 1990s. Further contracts covering the next decade were signed in 1990. But the political climate has changed considerably since then, and a question mark now hangs over the use of plutonium to generate nuclear power.

In addition, German utilities are no longer obliged by law to reprocess their waste. Last year the auditor-general's office reported that it cost twice as much to reprocess spent fuel as it would to place it into long-term storage, concluding that utilities were being unfairly required to operate uneconomically.

In response to this report, to pressure from environmentalists, and a desire to bring Germany into line with other EU countries, the nuclear law was amended in June to give utilities a choice in how they handle their waste. As a result, they are considering pulling out of their reprocessing contracts.

Withdrawal would, however, incur serious penalties, and these get stiffer each year. For example, if German utilities were to pull out of their contracts with Cogema immediately, it would cost them 10–15 per cent of the total 10-year bill. If they wait five years, they would become liable for 95 per cent of total costs.

The contracts with BNFL are believed to include similar penalties, but as the post-2000 contracts start two years later than Cogema's, there is less urgency.

Preussenelektra, Germany's biggest nuclear utility, wants Cogema to renegotiate more favourable terms. Other companies simply want to cancel their contracts, calculating that if all German companies did this immediately, paying instead for direct dis-

posal, they would make a total saving of up to DM5 billion (US\$3.15 billion).

In response to this pressure, Cogema has tried to increase its attractiveness to the utilities by offering solutions to Germany's perennial political problem of radioactive waste storage.

A leaked document revealed that, at least for Preussenelektra, Cogema has proposed three options: to store spent fuel without obligation to reprocess for up to 15 years, to reprocess it and store the plutonium on French soil, or to reprocess it and pass the plutonium on to a "third party".

But under French law radioactive waste may be stored only as long as is technically required for reprocessing. It must then be returned to its country of origin.

Cogema admits that the proposals were being discussed, and that they would not be permitted under this law — but defends its proposal on the grounds that nothing has been signed. BNFL is not bound by a similar law, and can therefore offer longer-term storage of either nuclear waste or separated plutonium.

Allison Abbott

. . . as hopes rise for German MOX factory

Munich. The German engineering company Siemens won an important victory last week when the highest administrative court in Germany overturned previous judgements by lower courts, and ruled that the company can continue construction work on a planned mixed-oxide (MOX) nuclear fuel plant in Hanau near Frankfurt.

At the same time, the opposition Social Democrat Party (SDP) has said that, if it takes office in October's general elections, it is willing to increase its support for renewable energy research conducted by the company to compensate for the promised banning of MOX production.

A spokesperson for Siemens said that the court's ruling would encourage the company to continue its work on the new plant. The plant was originally planned to start operating by 1992, but has been subject to continual delays because of objections from environmentalists.

But the battle is far from over. Challenges to six building and operating permits are awaiting court decisions. And if Germany's political complexion changes after the federal elections in October, the plant is likely to be closed permanently.

The original construction and operating licences were granted to Siemens by Hessen's previous government, led by the Christian Democrats. But when the SDP took office in 1991 in coalition with the

India's spending on R & D falls far behind promises

New Delhi. India's national expenditure on research and development (R&D) as a percentage of the gross national product (GNP) continued to fall during 1992–93, according to data published by the Indian Department of Science and Technology (DST).

The country's total expenditure on R&D was US\$1,754 million in 1992–93. In absolute terms, this represents an increase on the previous year. But as a fraction of GNP, the total budget has been declining steadily since 1988–89, when it peaked at 0.96 of GNP per cent.

In 1991–92, this fell to 0.84 per cent of GNP, and it dropped still further to 0.83 per cent during 1992–93. This is far below the proposed target of two per cent contained in a new science and technology policy awaiting approval by the cabinet.

According to the DST report, most of this money still comes from central and state governments. The private-sector share of national R&D expenditure was only 15 per cent of the total, and industry spent only 0.57 per cent of its sales turnover on R&D.

K. S. Jayaraman

Greens, each individual licence was challenged, primarily on safety grounds.

In addition, the newly installed Hessen government stopped the original MOX plant in Hanau from operating, and Siemens finally decided to close the plant in April this year (see *Nature* 369, 6; 1994).

German power utilities are now negotiating supplies of MOX — fuel elements that contain plutonium and uranium oxides — from British Nuclear Fuels (BNFL) in the United Kingdom and Cogema in France. Many of Germany's conventional uranium-burning power stations have been adapted to burn MOX to use up the plutonium they produce as waste.

At present, Siemens has the active support of the federal government in Bonn, a coalition led by the Christian Democrats. But the situation will be different if the SDP wins October's elections, as the party is committed to banning the commercial use of plutonium, including the production of MOX.

Siemens has another contingency plan. It has been discussing the possibility of selling equipment from the Hanau plant to BNFL, which is already planning to build a commercial MOX plant at Sellafield in northwest England. Helmut Rupa, deputy director of the Hanau factory, says that Siemens will reopen these talks in the autumn if the threat from the SDP materializes.

Allison Abbott

Bid to protect wolves from genetic pollution

Paris. A group of European animal rights and conservation organizations announced last week that it is setting up a European 'Wolf Federation' to seek better legal protection for the wolf, an endangered species in most parts of Europe — both physically and by 'genome pollution' through contact with both feral and domestic dogs.

According to the federation, protection is patchy and poorly enforced, while wolf populations are increasingly threatened by hunting and the destruction of their habitats. Healthy but declining groups still exist in parts of Eastern Europe, Russia and Siberia. But most groups elsewhere in Europe are small and scattered — Italy, for example, has fewer than a hundred animals.

Josh Ginsberg, a researcher at the Institute of Zoology at London Zoo, points out that in addition to its traditional enemy, man, the wolf is now under threat from domestic dogs, which have been able to exploit the reduced size of the wolf population to mate successfully with females.

Researchers have recently discovered that supposedly purebred European wolf populations *Canis lupus lupus*, a subspecies of the gray wolf *Canis lupus*, are in fact mainly hybrids between wolves and dogs.

In order to tackle the problem, researchers at London Zoo and elsewhere are seeking to establish new captive populations of the European wolf. According to Douglas Richardson, assistant curator of mammals at London Zoo, this entails identifying purebred individuals using genetic fingerprinting and other techniques, and then establishing stable founder populations.

Most wolves now kept in British zoos are American hybrids. Richardson says that the first aim is to replace these with European wolves, adding that captive breeding pro-

grammes could be used either to top up existing populations or to reintroduce wolves into areas from which they have disappeared.

The British organization Wolfwatch is working with London Zoo on the possibility of reintroducing European wolves into Scotland: with more than 300,000 red deer there is more than enough food as well as space to support a significant wolf population.

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

A dying breed? Efforts are needed to preserve the European wolf's genetic identity.

But Ginsberg admits that such a move is "politically unlikely". Hunters and farmers are likely to object, he says, while the public still regard wolves as dangerous, and the threat from dogs remains high. "Little Red Riding Hood has a lot to answer for," says Richardson.

But while zoologists and public associations are both backing wolf conservation, cooperation is often hampered by mutual distrust. Many of the groups making up the new federation say they oppose keeping wolves captive in zoos, and refuse to accept that zoos have a role to play in conservation.

For their part, zoologists often accuse amateur conservation groups of failing to take scientific advice into account, and of squandering resources on highly publicized but ineffective conservation schemes.

Richardson points out, for example, that the genetic make-up of founder populations needs to be monitored carefully, starting with captive groups. He also claims that amateur groups often fail to ensure that the reasons for the extinction of a species have been removed before attempting to reintroduce it. "They have to come to us if they want expertise," he says.

One event helping the European wolf to make a comeback is the removal of the fences and landmines in Eastern and Central Europe following the end of the Cold War. As a result, wolves have begun to spread from Poland into Germany, for example, says Richardson. The large-scale attempts to eradicate rabies from the mainland of Europe using a genetically engineered vaccine should also assist in conservation schemes.

Declan Butler

Italy promises appointment reforms

Munich. Stefano Podestà, Italy's minister for research and universities, has taken up a proposal originally made by Umberto Eco, professor of semiotics at the University of Bologna and author of *The Name of the Rose*, for reforming the country's controversial system of academic appointments.

The proposed system would involve drawing up a list from which universities could select individuals to fill vacant jobs. Podestà hopes that this and other suggestions will both increase the efficiency of the appointments system and reduce alleged malpractice.

At present, all academic posts in Italy are filled by national competitions (*concorsi*), which are supposed to take place every five years. National committees, one for each subject, select successful applicants who are then allocated to a university, not necessarily of their choice.

Italian scientists complain that the real selection of candidates results from behind-the-scenes dealings that are not always based solely on scientific merit, but can include deals between powerful faculty heads.

Addressing the final cabinet meeting before the summer recess, Podestà proposed

that committees should select more candidates than there are places, creating a national list from which universities would choose candidates when a vacancy arises, instead of having to wait for the next *concorso*.

Podestà said that he wants to simplify the procedure for electing the selection committees, and to give his ministry the right to appoint foreign academics to these committees "as necessary". He also outlined proposals to abolish the post of associate professor, leaving only two academic ranks: junior researchers (who at present do no teaching) and full professors.

While some have welcomed the change, Carlo Bullettì, secretary of the Italian Federation of University Researchers, which has long campaigned for a change in the academic promotion law, says it is too early to judge whether the proposals really offer an improvement on the current system.

Particular concern has been expressed that if the position of associate professor is eliminated, Italy would be out of line with most other countries. Podestà will draw up a more detailed parliamentary discussion paper in September.

Alison Abbott

Stars in the eyes

Washington. Although the world's space-faring nations are hard pressed to launch even a modest-sized space station into Earth orbit, there's no harm in thinking grander thoughts. So about a hundred scientists, engineers and visionaries will gather in New York on 29 August to discuss the topic: 'Practical Robotic Interstellar Flight: Are We Ready?'

Under consideration will be where to go (possible planets around other stars, the Oort cloud), and how to get there (everything from fusion rockets to magnetic sails). The four-day conference is sponsored by the Planetary Society, a space advocacy group, with the British Interplanetary Society and several other organizations. □

More funds urged for contraceptive research

London. The United Nations Population Fund (UNFPA) is calling for more effort and money to be spent on research into new methods of contraception as an essential part of stabilizing global population growth.

In its latest *State of World Population* report, published this week, the UNFPA highlights in particular the need for the development of more effective methods of male contraception, and of female methods that offer protection against both pregnancy and sexually transmitted diseases (STDs), as two of its most urgent tasks.

Other needs listed in the report included recognizing the position of reproductive health care and family planning services as essential parts of primary health care.

The UNFPA proposes that 20 per cent of national resources and international aid should be earmarked for human development, estimating that funding "will have to be raised considerably" to a minimum of US\$17 billion in 2000.

The report has been produced in advance of the fifth International Conference on Population and Development (ICPD), to be held

in Cairo, Egypt, on 5–13 September.

It points out that last year, according to figures produced by the US Program for Appropriate Technology in Health (PATH), only around three per cent of the total value of global contraceptive sales was spent on contraceptive research and development, in contrast to the 16–19 per cent of revenue spent by pharmaceutical companies.

PATH, which is based in Seattle, Washington D. C., has pointed out that scientists are currently working on more than 50 potential new contraceptive products. But a lack of financial investment means that most of these are merely variations on existing methods of contraception.

UNFPA's proposals coincide with recommendations made at the International Symposium on Contraceptive Research and Development for the year 2000 and Beyond, held in Mexico City in March 1993. They stress in particular that special emphasis in contraceptive development needs to be placed on "women's perceived needs and priorities"; these include methods that are under the user's control and that will protect

against STDs, and safe male methods.

These sentiments are echoed in the draft Programme of Action which participants hope will be agreed upon by member states of the United Nations at next month's conference in Cairo.

The draft calls for "a significant increase" in support from governments and industry to bring potential new methods into use. Areas singled out for attention are male and female barrier methods that both prevent pregnancy and protect against STDs, including HIV/AIDS.

But the chapter of the draft covering sexual and reproductive health research remains riddled, like the rest of the document, with square parentheses around such key phrases as 'fertility regulation' and 'sexual and reproductive health', indicating continuing opposition from the Vatican to attempts to make its scope as broad as possible. The words and phrases were placed in brackets after a failure to reach consensus on meanings or definition during the drawing up of the draft document at the third preparatory committee in New York in April 1994.

The Vatican is opposed to any suggestion that appears to support abortion and its attempts to water down the terms may overshadow much of the other work of the conference. Several participating countries — including the Philippines, Argentina, Chile, Iran — have already pledged their support to the Vatican.

Supporters of the programme prefer to stress the recent shifts in emphasis in population policies away from the consideration of population numbers in isolation towards population-related topics embedded in broader development policies.

Decisions in Cairo will be made on a simple majority vote, with more than 140 countries eligible to vote. But supporters of the draft fear that the necessary majority may only be possible with a 'watered down' final declaration. "I would be very disappointed and sad if that was so, and I confidently hope that common sense will prevail," says Anne McLaren, foreign secretary of the Royal Society, which was one of the main organizers of an international 'science summit' of academies held in New Delhi last year. The meeting resulted in a statement to achieve zero population growth (see *Nature* 366, 3; 1993).

The Inter-Academies Panel on Population Development, set up after the summit, will be represented in Cairo by Prakash Tandon, former president of the Indian National Science Academy. Tandon will present to the conference the statement agreed in New Delhi, which spoke of the need for an "integrated policy on population and sustainable development". McLaren describes the statement as "very forthright".

Maggie Verrall

'Population lab' urged for studies

London. A leading British surgeon is urging the United Nations to set up a "population control laboratory" to bring together scientific experts to tackle the problem of population growth. The proposal is made by Sir Roy Calne, professor of surgery at the University of Cambridge, and one of the world's leading transplant pioneers, in a book to be published next week.

Figures announced last week by the World Bank predict that the global population will increase from 5.6 billion now to around 8.5 billion in 2030, with almost 90 per cent of the increase in developing countries. Calne claims that population control is now "the most urgent task" facing the United Nations.

Calne argues in his book, *Too Many People*, that advances in science and technology are partly to blame. "If our species had not developed science, we would have muddled along," he says, pointing to trends such as rising global life expectancy, falling infant mortality, advances in tackling disease and other benefits of science and technology.

But science must now play a role in controlling reproductive rates, says Calne,

who wants a non-political, non-religious centre supported by the UN. He suggests that one step in this direction would be to establish a population control laboratory, run on a "non-hierarchical model" under the auspices of the UN.

Calne proposes that such a centre be located somewhere such as New Delhi in India where the problems of population growth are most in evidence. He claims that the idea already has latent support among scientists, pointing out that two years ago, for example, the Union of Concerned Scientists published a document on the problems of population growth, *Warning to Humanity*, signed by 1,600 scientists.

The suggestion has received a mixed reception among the scientific community. Robert May, for example, Royal Society research professor in the department of zoology at the University of Oxford, says he thinks it is an "excellent idea"; but, he goes on to add, given the reality of international organizations, "I don't hold out much hope for it".

Anne McLaren, foreign secretary of the Royal Society, points out that the World Health Organization already runs research programmes in the countries where there is the greatest need, which is "much better than the thought of some great centre that's going to try to impose solutions". Calne appears unfazed by such criticism. "It seems to me the problem is like a runaway cart and we must use scientific methods to pull on the reins," he says.

M. V.



Calne: 'science must give answers'.

Cancer researchers say NHS reforms undermining trials

London. The British government has been warned that its attempts to reform the National Health Service (NHS) by reducing the role of central government in the allocation of resources is seriously undermining clinical trials of new cancer treatments.

The latest in a series of warnings about the impact of the NHS reforms on biomedical research has come from the United Kingdom Coordinating Committee on Cancer Research (UKCCCR), a body representing the major organizations funding cancer research in the United Kingdom.

Its conclusions reinforce those of a report, commissioned by the government from a group headed by Anthony Culyer, professor of medical economics at the University of York, which are believed to confirm that Britain's biomedical research base is being seriously threatened by the reforms.

Although delivered to ministers in April, the Culyer report has not yet been published — partly, it is widely believed, because of the embarrassment its conclusions are likely to cause the government, as they run contrary to its efforts to establish an internal marketplace for all NHS expenditure (see *Nature* 369, 514; 1994).

John Smyth, professor of clinical oncology at the University of Edinburgh, and chairman of the UKCCCR, says his committee's main concern is that government attempts to reduce health service costs by increasing competition for funds are "against the cooperative spirit of research" needed to carry out large-scale clinical trials.

He also claims that the attempts to address shortcomings in research support by making additional money available to teaching hospitals under the so-called Service Increment for Teaching and Research (SIFTR) scheme have failed to address the problem adequately, as it is often difficult to tell what such funds are being used for. "SIFTR is not the answer," says Smyth.

A survey conducted by a working party set up by the committee has revealed that more than half of those conducting clinical trials of new cancer treatments say they are experiencing difficulty in sustaining their involvement in such research. Factors leading to reduced effort include a lack of time and of support staff, with many respondents saying that the time now spent in administrative and managerial committees means that less time is available for research.

In its report, published in last week's issue of the *British Medical Journal*, the working party says that "the financial imperative from a market-based NHS to avoid the 'unnecessary' expense in undertaking a [clinical] trial is the basis of real conflict".

The working group suggests that one

remedy would be to "top-slice" funds allocated to local health authorities, and develop a mechanism for distributing these to support national clinical trials. It also calls for a "radical change" in the way in which SIFTR funds are handled, allowing for a greater proportion to be allocated to following patients involved in such trials.

The report's main conclusions have been endorsed by individual cancer research organizations, concerned that reduced funding for patient care is requiring them to carry an increasing share of the costs of carrying out trials of new drugs and therapies.

But the government has so far given little indication that it is prepared to acknowledge that its NHS reforms have created a major problem for research. John Bowis, parliamentary secretary at the Department of Health, said in a statement that he rejects the UKCCCR's view that changes in the NHS signal a reduction in the government's commitment to ensuring that world-class clinical research continues to thrive.

Bowis pointed out that the government has already set a goal of increasing NHS spending on research and development to 1.5 per cent of its total expenditure. But he pointed out that the decision to ask for recommendations for "improvements" from Culyer had reflected the government's acceptance that there is a need to look again at mechanisms intended to ensure that such research and development is being properly funded. He also promised that the Culyer report will be published "shortly".

Such statements have so far failed to reassure the government's critics. "Our main concern is that its current attitude is generating a growing apathy among clinical researchers, many of whom are now seeking research opportunities abroad," says Smyth.

Such fears have recently been reinforced by the decision of several leading biomedical researchers — including Bob Williamson, professor of biochemistry at St Mary's Hospital and a pioneer of cystic fibrosis treatment, and Lucio Luzzatto, former head of haematology at the Hammersmith Hospital, both in London — to leave the country.

In announcing his planned departure last month, Williamson said that one of the key factors leading to his decision had been the growing difficulties of obtaining adequate support for research under the new organization of the NHS. The issue is already being taken up by the opposition Labour Party, which claims that the government's commitment to the 'internal market' policies for the health service "is effectively pulling the plug on research and development".

David Dickson

Conference trail can leave behind a good image

London. Been to any good conferences lately? If not, then it may be time to make amends. A report published in the journal *Social Studies of Science* (24, 513–548; 1994) shows how an individual's record of attendance at scientific meetings may be as good a way of assessing intellectual quality as an analysis of the impact of his or her scientific publications.

Thomas Söderqvist of Roskilde University in Denmark and Arthur M. Silverstein of Johns Hopkins University School of Medicine in Baltimore, Maryland, analysed the attendance of 4,806 delegates who took part in 88 immunology meetings between 1951 and 1972, a period during which immunology became established as a distinct discipline.

The majority of participants (72 per cent of the total) attended only one meeting during the 23-year period. But almost all of the 1.6 per cent (79) who went to ten or more meetings can be identified by



other criteria as leaders in the field.

Significantly, these criteria extend beyond simple citation records; some of the most frequent attendees had not made important discoveries themselves, but had been noted for (and made their contribution to the field through) their ability to spot talent in others.

Söderqvist and Silverstein's main goal was to spot trends in the structure and dynamics of an emerging field. But the study also seems likely to attract interest as a way of reinforcing other techniques for judging scientific quality.

The two researchers describe it as "astonishing" that analysis of the participation of scientists at meetings — a "pervasive and integral part of science" since the seventeenth century — has been generally neglected as a way of identifying research trends.

Henry Gee

Biodiversity in Switzerland

SIR — The Swiss National Science Foundation (NF) recently funded a large (US\$3.5 million) biodiversity research programme. Oliver Klaffke, in the News article "Ecologists clash over too academic research" (*Nature* 368, 283; 1994), found that some ecologists from non-academic organizations in Switzerland were unhappy with the funding process for this programme and upset that most of the funds went to academic institutions for basic research. The article did not, however, discuss the original aims of the biodiversity programme, the scope of the current research or the relationship between basic and applied research. As co-authors or participants in five of the projects funded by the NF, we would like offer our views.

The initial call for proposals from the NF stated that the objectives of the biodiversity research programme were (1) to investigate the origins of biodiversity, especially of genetic diversity; (2) to study the dynamics of populations and the human influences on them; (3) to study the value of biodiversity; and (4) to evaluate strategies for the long-term conservation of biodiversity. Interdisciplinary, coordinated programmes were especially welcomed.

Twelve of the 20 funded projects were collectively designed by a 'diverse' group of scientists, mainly at the University of Basel, who developed a coordinated research programme to provide insights into the nature, maintenance and values of biodiversity ranging from the gene to ecosystem levels. The Basel scientists have chosen to concentrate their research at common grassland sites in the Jura Mountains, to share experimental infrastructure, and actively to exchange data. This group of scientists includes animal ecologists, biogeochemists, ecosystem modellers, evolutionary geneticists, meteorologists, plant population biologists and plant pathologists, as well as a lawyer concerned with legal means for the conservation of biodiversity. The two largest experiments in this programme seek to understand how two important human induced disturbances, habitat fragmentation and global change, will affect biodiversity. The goal of the habitat fragmentation study is to determine how genotypic and species diversity are affected by small-scale, experimental fragmentation in a species-rich, limestone grassland. These grasslands, like many other complex ecosystems, are becoming smaller and more isolated because of land use changes. Failure to understand the consequences of this fragmentation may lead to conservation strategies that fail to maintain biological diversity. The biodiversity/CO₂ study will help (1) to

determine the role of biodiversity in ecosystem function and (2) to determine whether the future increases in atmospheric CO₂ concentration will cause gene shifts within species, alter species dominance or lead to species extinctions. This study has been selected by an international board to be a core research project in the Global Change and Terrestrial Ecosystems programme of the International Geosphere-Biosphere Programme

(*Ambio* 23, 74–76; 1994). The programme also includes studies of the generation of genetic diversity using microorganisms and the population dynamics of rare plant species. We firmly believe that such an integrated, experimental approach to understanding biodiversity is not "too academic". It is a prerequisite for successful management of biodiversity and for assessing its value.

John A. Arnone III
Plant Ecophysiology,
Michel Blot
Evolutionary Genetics,
Biozentrum,
Paul Leadley,
System Ecology,
Diethart Matthies
Plant Population Biology,
Ian Sanders
Plant and Fungal Ecology,
University of Basel,
CH-4056 Basel, Switzerland

What next?

SIR — What next — is *Nature* going to support an obligatory licence to ingest butter, bacon, beer or other allegedly harmful substances subject to positive medical or other proof by licensee that the substances have not yet caused harm? Is *Nature* suggesting enforcement through permits and licences of recommended intake of calories, fat, sugar and salt? Should people be force-fed optimal food-tablets and otherwise restricted to reduce their perceiving personal risks?

Are prohibition, rationing and/or licensing of "tobacco (cannabis)" and the like, which mainly cause harm to voluntary consumers of them, *Nature's* choice of tools to regulate the quality of life of intelligent citizens in free societies? And is *Nature* contemplating sanctions (the severity of which I dare not consider) for the wrong-doers?

Knowing *Nature*, I would be surprised, but reading "When will cigarettes be smoked out?" (*Nature* 369, 691–2; 1994), one can but wonder.

Christian Hansen
Østerbrogade 118, 3. tv,
DK-2100 Copenhagen,
Denmark

Backward step

SIR — The 1994 meeting of the British Association is nearly upon us.

It is for the advancement of science, a subject that involves many facts as well as methods, hypotheses, concepts and opinions, a subject in which newly discovered results and ideas usually need to be related to old ones in order to be understood. For example, it makes little sense to seek to explain what a Higgs boson is, if not the sociology of why we should search for one, to someone who has only a hazy notion of an atom and no notion of an electron.

These thoughts prompt the suggestion of two *gedanken* experiments. Let all the lectures at the 1993 British Association meeting, except those by Messrs Lawson and Waldegrave, be repeated at the 1994 meeting. How many of those attending would be aware of the repetition?

Another experiment would be to examine orally all those who have read last year's explanations of the Higgs boson. What would their average grade be?

I am pessimistic about the results of these experiments because the dilemma of making science widely intelligible and widely attractive is so difficult to resolve. It does science and the nation a disservice to underestimate the difficulty of resolving the dilemma.

Phillip Drazin
University of Bristol,
School of Mathematics,
University Walk, Bristol BS8 1TW, UK

Cancer trials

SIR — In your News story on the breast cancer prevention trials (*Nature* 369, 516; 1994) you state that the Imperial Cancer Research Fund (ICRF) is recruiting women for the UK trial. In fact the UK trial is run by the United Kingdom Coordinating Committee on Cancer Research and women are recruited by the trial coordinator and by the participating centre, not the ICRF. The ICRF's role is, together with the Cancer Research Campaign, to provide the funding for the trial.

Jean Mossman
United Kingdom Co-ordinating
Committee on Cancer Research,
PO Box 123, Lincoln's Inn Fields,
London WC2A 3PX, UK

Correction

SIR — In "Delaney and cancer" (*Nature* 368, 580; 1994), lines 41–42 should read " 9×10^{18} molecules" as originally published in reference 3, rather than 9×10^{10} .

Thomas H. Jukes
Department of Integrative Biology,
University of California, Berkeley,
Oakland, California 94008, USA

Primary care is not the answer

The idea that the United States can solve major problems in medical care by mandating that in the future 50 per cent of physicians go into primary care is wishful thinking.

PRESIDENT Bill Clinton's heroic hope of winning a place in history as the visionary who made good medicine instantly available to every man, woman, and child in the United States (while reducing the deficit at the same time) is likely to be dashed by a contrary Congress that is consumed by disagreement on health legislation. That is a good thing, not because Clinton's vision is wrong but because many of the details in his health-care bill are so poorly thought out as to be dangerous. (It should also be said clearly that the real impetus to reforming the health insurance system in the United States is cost, not humanitarianism.)

The big arguments in Congress are about whether all employers (including small businesses) should be required to pay insurance premiums and whether new coverage can be extended to the whole of the population right away.

But less visible provisions in the various versions of health legislation are (or ought to be) equally contentious. One is a provision to force 50 per cent of physicians into what is called 'primary care', known once upon a time as general practice. During the past two decades, the United States has alternately concluded that it had too few physicians, too many, and now, that it has a surplus not of physicians generally but of specialists in particular. From start to finish, the data have been weak but have nonetheless served as the impetus to legislative remedy. So today the fashion is to talk about an oversupply of specialists and the need to create by 2000 (or so) what is facetiously called 'the 50 per cent solution'.

The premise is quite simple. An analysis of the physician population shows a predominance of relatively expensive specialists (cardiologists, gastroenterologists and oncologists, for instance) in cities and suburban areas, and a deficiency of primary-care doctors to provide vaccinations and be there to catch minor illness before it becomes serious disease.

C. Everett Koop, the former US surgeon general known for his outspoken pronouncements during the Reagan administration, recently argued the case to the general public through an article in the 3 July issue of *Parade*, a tabloid supplement published in nearly every Sunday paper in the country. Koop noted that "in Britain, 72% of physicians are primary-care doctors. Canada has 54%. Germany and France have 47% and 48%. We have 29% and the number is falling." *Ergo*, the United States needs more

primary-care doctors too. Speaking eloquently of the virtues of primary care, Koop aimed at the heart: "The patient is more than the aches and pains he or she brings to you. You may come in to see the doctor about a pain in your elbow, but you may have problems that are economical, emotional, psychological or spiritual." The elbow, Koop argues, is the opening for the primary-care doctor (cum financial adviser, psychiatrist and cleric) to solve problems of all sorts. In Koop's ideal world the patient will go home saying "I don't have a problem with my elbow any more." Why? Because the doctor has changed his attitudes about life."

This is patent nonsense and dangerous besides. It perpetuates (in fact, reincarnates) the myth of physician as God, and ignores the possibility that the problem with the patient's elbow is rheumatic or orthopaedic in origin and has nothing to do with his or her bank account or soul. It fosters the myth that real people with nice, friendly general practitioners do not get really sick. That makes modern medicine the enemy.

An equally interesting new account of medicine in Britain makes the countervailing argument, although the information has not yet been part of the US debate. According to recent data, people in Britain are more likely to die of cancer (particularly lung, rectal, prostate, stomach and breast cancer) than their counterparts in the United States and Europe. Why? If Kenneth Calman, director of the UK Department of Health, is right, Britain has too few oncologists. Only six out of every 10 women diagnosed with breast cancer in Britain will be alive five years later. In the United States, eight of those same 10 women will survive in relatively good health for more than five years.

There are two important points here. First, the debate about primary-care physicians versus specialists is far more complex than has been portrayed. No one can argue persuasively that modern science has not led to new and useful therapies that are beyond the skill of the generalist, any more than one can argue that all doctors should be highly technically skilled specialists. Furthermore, in the United States at least, there are no data (or provisions in legislation) that will demonstrably solve the problem of physician distribution in poor and rural areas simply by mandating that more medical students should go into primary care.

Second, and even more important, the US education system has, for a quarter of a

century or more, been geared to producing specialists — especially at the major academic medical centres in the United States. It is these people (and this system broadly defined) that are the core of the research enterprise that has done so well in advancing not only the care of patients with complex illnesses but also understanding of basic biological phenomena that will produce the next wave of medical progress.

It may be wise to change the system, and the numbers and kinds of physicians who are educated. But not for the reasons put forward so far. The point that health-care reform is really about money, not about high-quality medical care, is worth reiterating because it has become, in part, the basis for the arguments about how many physicians are needed in this or that speciality. The fact that the United States has more specialists than other countries is not, *per se*, an argument for fewer specialists. Nor is the argument, frequently cited in the medical literature, that managed care systems such as Health Maintenance Organizations (HMOs) employ more generalists than specialists a sensible one on which to base legislative decisions. That argument assumes that HMOs do an equally good job of caring for or referring patients with complex disease — an assumption that has yet to be proved.

A couple of facts need to be faced. First, modern medicine — born of research — saves lives but does not always save money, if for no other reason than that the patient cured of one disease will live to get another. Second, it is not always true that medicine in other countries equals that in the United States. Rather, people elsewhere seem to accept levels of care and limits to access that Americans (at least for now) will not. It is said that Americans believe that death is an option. Hyperbole aside, they certainly believe that premature death — death that can be stayed by the hand of medicine to give an individual a longer and reasonably good life — should not be an option. (We are not talking here about the heartlessness of putting the terminally ill on respirators so that they can live out their last days in personal anguish at high cost to society.)

But for appropriate life-prolonging care, specialists and modern biomedical research are essential and should not be ignored in the headlong rush to pass legislation that, in the end, may damage science and, therefore, medicine.

Barbara J. Culliton

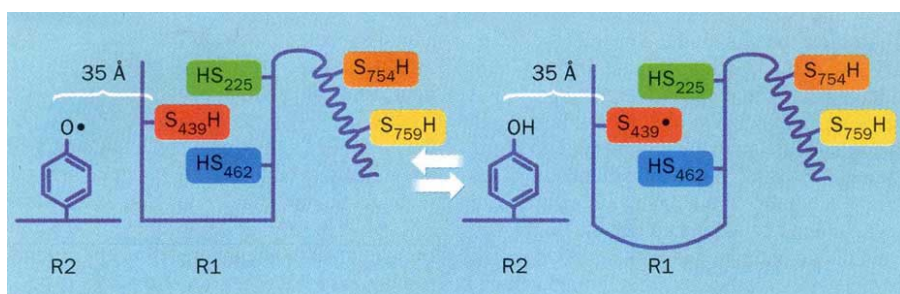
Controlling radical reactions

Joanne Stubbe

FREE radicals let loose in the cell can perpetrate all kinds of damage, including mutagenesis and molecular degradation that may manifest as disease. Could it be a coincidence that an important step in DNA biosynthesis, namely the conversion of ribonucleotides to deoxyribonucleotides by ribonucleotide reductases (RNRs), involves stable amino-acid radicals as well as transient amino-acid and nucleotide radical intermediates (ref. 1)?

and are themselves oxidized to a disulphide^{7,10}. A third cysteine at position 439 was postulated to be transiently converted into a thiyl radical which then initiates nucleotide reduction by abstracting the 3' hydrogen atom of the substrate⁸. This thiyl radical on R1 may be the partner in the electron transfer to the tyrosyl radical of R2.

The structure determined by Uhlin and Eklund⁴ positions cysteines 225 and 462



Escherichia coli RNR is composed of two subunits, R1 and R2. The unusual dinuclear iron centre and tyrosyl radical cofactor reside on R2. The binding site for the ribonucleotide substrates resides on R1, as do the five cysteines that are postulated to be essential for deoxyribonucleotide formation. Communication between the two subunits is essential for catalytic activity and is thought to involve long-range electron and proton transfers.

In the RNR enzyme from *Escherichia coli*, which has two subunits R1 and R2, the subunit R2 has a stable organic radical on a tyrosine residue adjacent to a dinuclear iron centre² which is essential for nucleotide reduction: a one-electron reduction of the tyrosyl radical to tyrosine destroys all catalytic activity. The stability of this tyrosyl radical in subunit R2 is evidenced by its half-life of days — tyrosyl radicals in solution last only for a matter of microseconds. The actual nucleotide reduction has been proposed to occur on the R1 subunit (the 'business' end of RNR), with the R2 tyrosyl radical acting as a radical chain initiator and communication between subunits being effected by an electron transfer process (see figure)^{1,3}. Now this hypothesis is given added weight by the newly determined X-ray crystallographic structure at 2.5 Å resolution of the R1 subunit from *E. coli*, described on page 533 of this issue⁴.

The RNR from *E. coli* is a prototype of the mammalian and the herpes viral enzymes⁵ and was the first protein found to contain a stable organic radical². Experiments comparing the action of wild-type RNR enzyme on substrate analogues (mechanism-based inhibitors)⁵ and using site-directed mutant protein analogues⁶⁻⁹ have indicated that there are three cysteines located in the active site of R1. Two cysteines at positions 225 and 462 directly provide the reducing equivalents required for nucleoside diphosphate reduction

within 6 Å of a third cysteine at residue 439. These cysteines are detected in the structure as a disulphide, which is consistent with the role just discussed, fulfilling the expectations of biochemical studies. Also, the structure reveals that glutamate 441, not previously identified, is poised to participate as a general acid/base catalyst in the mechanism of reduction.

Unfortunately, neither substrate nor allosteric effector have been co-crystallized in the reported R1 structure. The conformational flexibility of this protein, demonstrated by its ability to reduce purines and pyrimidines only in the presence of the appropriate allosteric effector, suggests that additional groups may be important in catalysis after structural reorganization.

Two additional cysteines have been proposed to play a part in nucleotide reduction, by transferring reducing equivalents provided by the protein reductant thioredoxin (or glutaredoxin) into the active-site disulphide by disulphide interchange (see figure)^{6,9,11}. These cysteines (at positions 754 and 759) are not resolved in the X-ray structure because they are in the C-terminal 24 amino acids of R1, which are thermally labile. The flexibility of this tail provides a reasonable explanation for how dithiothreitol, a small organic reductant, can bypass the thioredoxin redox shuttle and provide reducing equivalents directly to the active-site cysteines at 225 and 462. The structure of R1

therefore nicely complements earlier results and offers clues through the discovery of glutamate 441 and of the flexibility of the C-terminal cysteines as to how this complex protein might function.

How has this structure affected our thinking on the nucleotide reduction process? As indicated in the figure, the tyrosyl radical on the R2 subunit of *E. coli* RNR was presumed to generate a thiyl radical at cysteine 439 on R1 as a result of long-range electron transfer. The X-ray structure, however, suggests that cysteine 439 is protonated within the active site. Its environment does not appear to have residues at a position or secondary structure appropriate to perturb its pK_a, which is necessary to generate a thiolate essential for electron transfer. The structure therefore suggests that an alternative mechanism should be considered for generating the thiyl radical: hydrogen atom abstraction (or the equivalent, a coupled electron and proton transfer). The structure of R1 and docking experiments with R2 suggested to the authors⁴ that tyrosines 730 and 731 play a key role in this communication, over a distance of some 35 Å, between the active site of R1 and the tyrosyl radical on R2, a pathway that might involve electron hopping between aromatic amino-acid residues.

This structure will have an important impact in three ways: first, the confirmation that cysteine 439 resides in the active site of R1 will stimulate intense effort to demonstrate the existence of transient thiyl radicals, which are unprecedented in catalysis; second, the tentative location of one of the allosteric binding sites adjacent to the active site will encourage investigation of the basis of substrate specificity and the role of allosteric effectors; third, the structure will help clarify how the two subunits communicate. The R1 structure adds an additional piece to the complex puzzle that should eventually explain how this amazing protein can tame highly reactive radical intermediates and keep control of the substrate specificity essential to the fidelity of DNA replication. □

Joanne Stubbe is in the Department of Chemistry, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA.

1. Stubbe, J. *J. Biol. Chem.* **265**, 5329–5332 (1990).
2. Sjöberg, B.-M., Reichard, P., Gräslund, A. & Ehrenberg, A. *J. Biol. Chem.* **252**, 536–541 (1977).
3. Nordlund, P. & Eklund, H. *J. molec. Biol.* **232**, 123–164 (1993).
4. Uhlin, U. & Eklund, H. *Nature* **370**, 533–539 (1994).
5. Stubbe, J. *Adv. Enzym. relat. Areas molec. Biol.* **63**, 349–417 (1990).
6. Mao, S. S. *et al. Biochemistry* **31**, 9733–9743 (1992).
7. Mao, S. S. *et al. Biochemistry* **31**, 9744–9751 (1992).
8. Mao, S. S., Yu, G. X., Chalfoun, D. & Stubbe, J. *Biochemistry* **31**, 9752–9759 (1992).
9. Åberg, A. *et al. J. Biol. Chem.* **264**, 12249–12252 (1989).
10. Thelander, L. *J. Biol. Chem.* **249**, 2858–2862 (1974).
11. Lin, A. E., Ashley, G. W. & Stubbe, J. *Biochemistry* **26**, 6905–6909 (1987).

Kitty litter for carbon control

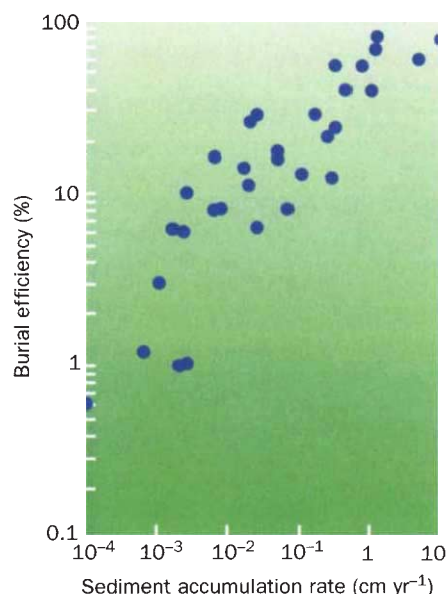
Cindy Lee

EVEN before the days when Darwin wrote about earthworms and soil humus, there was an interest in the degradation of organic carbon, which in nature occurs to a large extent through the action of bacteria and other microorganisms. Anyone who has taken out the garbage on a hot day is familiar with microbial decomposition of organic matter. Cat lovers know that cat litter or similar clay mineral products can adsorb and reduce the smell of organic matter and its degradation products. Other processes may be at work here also. Keil, Montluçon, Prah and Hedges¹ provide evidence on page 549 of this issue that the carbon adsorbed by such mineral matrices is protected from bacterial degradation and suggest that this way, labile carbon can be preserved in marine sediments.

Oceanographers have long been intrigued by the controls on carbon preservation because sedimentary organic matter is one of the chief stores of carbon on the Earth. Stored carbon can influence global atmospheric oxygen as net burial of organic matter produced during photosynthesis is balanced by an increase in free oxygen which increases the oxidation state of the oceans and atmosphere². Much of the organic carbon buried in sediments is deposited on the coastal margins, the type of site studied by Keil *et al.* What they find is, first, evidence that mineral sorption promotes carbon preservation. This has been shown before for various individual organic compounds. More important and somewhat surprising is that much of the organic matter that is preserved can apparently be desorbed and then degraded by microorganisms.

Mayer^{3,4} recently suggested that most of the organic matter in marine sediments is present in a monolayer coating on mineral surfaces. Incorporation into small pores may protect the organic matter from decomposition as pores smaller than about 8 nm would exclude entry by most of the hydrolytic enzymes necessary for decay, and larger pores could still inhibit enzyme function. Keil *et al.* find a strong correlation between organic matter and mineral surface area in a core from their coastal margin site, a perfect location to test the mineral protection hypothesis. They desorbed the organic matter from the sediment and subjected the resulting solution to degradation by aerobic bacteria taken from the overlying seawater column. Most of the desorbed organic matter degraded quickly — a remarkable result, because much of the adsorbed sedimentary organic matter, particularly in deeper sediments, has been considered refractory.

Protection by a mineral matrix is a sensible explanation for carbon preservation in soils and sediments, and helps explain the strong relationship between sedimentation rate and carbon burial efficiency that has been observed⁵ (see figure). Why then is there any controversy? Keil and colleagues carried out



The general trend of the relationship between carbon burial efficiency and sediment accumulation rate in the world's oceans. Burial efficiency is the proportion of the total carbon flux to the sediments that is buried or preserved below the zone of active decomposition. Data points are for locations representing a variety of depositional environments. (Figure adapted from ref. 5.)

their experiments with sediments where the correlation between mineral surface area and organic carbon was strong and almost constant. In sediments under the rich upwelling areas off California and Peru, there is more organic matter than can be accounted for by just mineral adsorption⁴. Mineral input to these areas is low, yet the organic matter content of the sediments is high and preservation is apparently enhanced. What is preserving organic matter in these sediments?

Historically, it has been argued that oxygen is a controlling factor. If anoxic decomposition of organic matter is intrinsically slower than oxic decomposition, more carbon will be preserved in the anoxic sediments deposited on coastal margins or other areas of high productivity. Others argue that high primary productivity in such areas results in a higher flux of carbon to the sea floor, so more carbon is indeed preserved, but not necessarily in proportion to the flux⁶.

Anoxia is a result of the decomposition of organic matter in these sediments but is not a controlling factor.

Canfield⁷ recently proposed a compromise. He argues that in areas of lower sedimentation that have low-oxygen or anoxic bottom waters, anaerobic organisms are unable to oxidize certain classes of compounds and carbon is preserved. But in anoxic areas of high sedimentation rates, burial efficiencies lie above the general trend shown in the figure and are decoupled from sedimentation rate. Protection by adsorption to minerals would be less important in these environments and so other preservation mechanisms must still be of consequence.

Where most of the organic matter is present on mineral surfaces, the finding that this matter is labile is a challenge to the preconceptions of many marine chemists. Traditionally, adsorption onto minerals is thought to promote condensation reactions, one of the first steps towards making the matter more refractory. Perhaps the results of Keil *et al.* give us an estimate of the initial rate of such processes, as the percentage of desorbable matter decreases with increasing age in their sample core.

Their study raises many other questions. Of prime importance is the source of the mineral material. If the minerals are land-derived, is the adsorbed terrestrial organic matter being brought to the sea labile, and what happens at the estuary interface? Is there an exchange of terrestrial and marine material during the particle flocculation and precipitation observed in these regions? Does mineral material of marine origin contain organic matter from the source organisms or has this been replaced with other matter? As dissolved organic carbon in sea water does not degrade under similar conditions, the fact that sediment-adsorbed organic matter is labile suggests that it may be material from the original organism that has never been desorbed. We will need to know far more about the composition of the desorbed matter and the mechanisms of adsorption to answer these questions.

Another puzzle is why, if the organic matter desorbs so easily, it has not decomposed. Organic matter binds to minerals to different degrees⁸. The techniques used by Keil *et al.* remove reversibly adsorbed organic matter, and as this

1. Keil, R. G., Montluçon, D. B., Prah, F. G. & Hedges, J. I. *Nature* **370**, 549–552 (1994).
2. Garrels, R. M. & Perry, E. A. Jr in *The Sea*, 303–336 (Wiley, New York, 1974).
3. Mayer, L. M. *Geochim. cosmochim. Acta* **58**, 1271–1284 (1994).
4. Mayer, L. M. *Chem. Geol.* **114**, 347–363 (1994).
5. Henrichs, S. M. & Reeburgh, W. S. *Geomicrobiol. J.* **5**, 191–237 (1987).
6. Pedersen, T. F. & Calvert, S. E. *Am. Assoc. Pet. Geol. Bull.* **74**, 454–466 (1990).
7. Canfield, D. E. *Chem. Geol.* **114**, 315–329 (1994).
8. Henrichs, S. M. & Sugai, S. F. *Geochim. cosmochim. Acta* **57**, 823–835 (1993).
9. Wang, X.-C. & Lee, C. *Mar. Chem.* **4**, 1–24 (1993).

desorbs from the surface in equilibrium with other dissolved material, it is susceptible to microbial attack and should decompose as quickly as free material. Why didn't it? In the deepest sediment, it had 500 years to desorb and degrade. If it is irreversibly adsorbed or desorbs only slowly as has been observed for some compounds, it should not have come off the mineral surfaces under the techniques employed.

The oxidation state of the material also plays a role in mineral adsorption. Metal oxides and other minerals in oxic sediments that are known to adsorb organic compounds can dissolve under anoxic conditions and release adsorbed organic

compounds that would then be subject to anaerobic decomposition⁹. It will be important to find out whether the adsorbed organic matter is as labile under anoxic conditions. In this respect, oxygen may play an additional role in preservation.

Where does this leave us? In short, 'kitty litter' control on carbon preservation is entirely plausible, yet it leaves many intriguing questions unanswered. □

Cindy Lee, of the Marine Sciences Research Center, SUNY, Stony Brook, New York 11794, USA, is currently at the Max Planck Institute for Marine Microbiology, Fahrenheitstraße 1, 28359 Bremen, Germany.

VERTEBRATE MORPHOLOGY

Return of the amphioxus

Henry Gee

TIME to dust off your textbooks and welcome back an old friend. After decades of frustration, work on the amphioxus — the closest invertebrate relative of the vertebrates — is starting to illuminate the problematic origins of our own particular corner of creation.

On page 563 of this issue¹, Garcia-Fernández and Holland describe a single collinear cluster of homeobox genes from the amphioxus *Branchiostoma floridae*, to match the four paralogous clusters known from mammals. Each of the amphioxus genes can be assigned to a vertebrate paralogy group, and they are arranged in the same order.

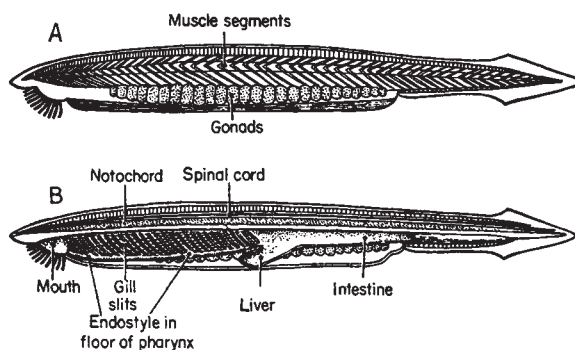
Of course, animals of all kinds have similar clusters, but the close relationship between amphioxus and vertebrates imbues this finding with a phylogenetic significance not lost on participants at a recent meeting*. Indeed, the large-scale duplication of the homeobox cluster and perhaps other genes might be related to the origin of those vertebrate structures that the amphioxus lacks.

Sans eyes, sans teeth, sans almost everything, the amphioxus looks like nothing more than a pallidly animated anchovy fillet. Its closeness to vertebrates is defined by its notochord, dorsal tubular nerve cord and paired somites; its deficiencies (particularly in the head region) throw the problem of the origin of the vertebrate head into sharp relief.

Most of the differences concern neural crest, a tissue which on its own, or through its interaction with mesenchyme, pro-

duces most of the structures we associate with the vertebrate head, such as paired eyes, bones of the face and jaws, teeth, craniofacial musculature and so on.

The amphioxus has none of this. Rather than stopping beneath the midbrain, as in vertebrates, the notochord runs right to



The classic textbook picture of the amphioxus, from a classic textbook, Romer's *The Vertebrate Body* (4th edition, W. B. Saunders, 1970). A, As seen through the transparent skin; B, a sagittal section.

the anterior tip; the cerebral vesicle — a negligible anterior swelling of the nerve cord — passes for a brain, terminating in front with a modest and solitary pigment spot; there are no teeth or other hard tissues, and the somites of the trunk run free to the front. An ancestor less gifted than this would be hard to find.

Not that this worries the amphioxus unduly: the two genera of amphioxus, *Branchiostoma* and *Epigonichthys* comprise about 25 species (S. G. Poss, Gulf Coast Research Laboratory, Mississippi, and H. T. Boschung, Univ. of Alabama). They are usually found half-buried in the sands of temperate and tropical inshore waters worldwide, where they strain algae through a complex pharyngeal filtration mechanism. Far from bathyal obscurity,

they can turn up in abundance in the most surprising places: the green, near-eutrophic Tampa Bay in Florida (especially the part close to the airport where a causeway runs across to Clearwater) is the amphioxus capital of the western hemisphere, yielding most of the animals used in the projects described here.

The headlessness of the amphioxus has set the agenda for a century of discussion of the 'head problem'. Is the vertebrate head an elaboration of the anterior region of an amphioxus-like condition, or did the evolution of the neural crest produce an entirely new structure² tacked on to the front end?

This problem has been abandoned by most experimental biologists as insoluble, yet at the turn of the century, morphology-inspired amphioxus research into just such questions assumed the kind of prominence that *Drosophila* work does today. So much so, that by 1932 Conklin³ prefaced yet another huge monograph with an apology to readers daunted by the prospect of reading even more on "such a well-worked subject as the embryology of Amphioxus". The decline set in when the possibilities for discovery offered by a light microscope had been exhausted. But new techniques (and a few old ones, used in new ways) have led to some real advances.

The ability to locate and map homeobox genes is one such technique. *Hox* genes in mice are never expressed more anteriorly than the rhombomeres of the hindbrain. Were the amphioxus the equivalent of a headless trunk, these expression patterns would run right to the anterior end. Not so: the expression domains of all the genes have clear anterior limits corresponding with various morphological 'landmarks' along the body^{4,5}, suggesting that a large part of the animal is equivalent to a head. So rather than being a completely new structure, the vertebrate head is, essentially, fancy neural-crest-derived bodywork welded to a plain and pre-existing chassis.

Other lines of evidence confirm this idea. Lacalli *et al.* reconstructed the neural architecture of the cerebral vesicle from electron micrographs of 5,000 serial transverse sections, digitized and fed into a computer to produce three-dimensional images (N. D. Holland, Scripps Institution of Oceanography, California). The result⁶ suggests that at least part of the neural tube can be regarded as an elongated hindbrain, with the addition of an eyespot believed (from its neuroanatomy) to be homologous to the paired eyes of vertebrates, which originate in the diencephalon. The amphioxus does not seem to have anything corresponding to a telencephalon or a midbrain.

*Fourth International Congress of Vertebrate Morphology, University of Chicago, 31 July – 4 August 1994.

GEOLOGY

Again, the closeness of amphioxus and vertebrates receives support from neuroanatomy (B. Fritsch, Creighton Univ., Omaha), myosin light-chain expression (D. A. Pace and L. Z. Holland, Scripps), DNA hybridization kinetics (G. N. Sideris, New York Univ.) and mitochondrial gene arrangements (W. M. Brown and L. L. Daehler, Univ. of Michigan, Ann Arbor). And scratch out the bit in your textbooks that says how the excretory organs of the amphioxus are like those in flatworms. This has been debunked by some neat histology showing that they are similar to vertebrate glomeruli, with some modifications allowing for their small size (E. E. Ruppert, Clemson Univ., South Carolina). Few of these advances would have been possible without modern technology, yet they lay some weighty problems to rest.

So too, though, does something as simple as a video film, showing how amphioxus larvae actually feed — a problem that had troubled the finest minds for a century (T. H. J. Gilmour, Univ. of Saskatchewan). Likewise, a 'timetable' of amphioxus development compiled from scanning electron micrographs (M. D. Stokes and N. D. Holland, Scripps) showing that the anus originates on the right, swinging over to the left later on (just like Van Wijhe said it did back in the 1890s, only nobody believed him). Submicroscopic detail of the amphioxus nervous system is being revealed through the ability to tag individual neurons with a fluorescent dye and trace their subsequent development (L. S. Demski and J. B. Morrill, Univ. of South Florida).

Yet the amphioxus still has many other old problems with which to challenge us. Still unsolved is the curiously lopsided development of the larva, in which precocious left gill slits appear long before the right-hand series. Whereas left and right somites in vertebrates are locked in phase with each other, amphioxus somites are prone to slippage with respect to one another and the notochord (R. Presley and T. J. Horder, Univ. of Wales, Cardiff, and J. Slipka, Charles Univ., Czech Republic). This could be one example of a tendency towards asymmetry seen in deuterostomes generally (R. P. S. Jefferies, Natural History Museum, London), which modern methods have so far proved powerless to explain. □

Henry Gee is an assistant editor of *Nature*.

1. Garcia-Fernández, J. & Holland, P. W. H. *Nature* **370**, 563–566 (1994).
2. Gans, C. & Northcutt, R. G. *Science* **220**, 268–274 (1983).
3. Conklin, E. J. *Morphol.* **54**, 69–151 (1932).
4. Holland, P. W. H., Holland, L. Z., Williams, N. A. & Holland, N. D. *Development* **116**, 653–661 (1992).
5. Holland, P. W. H., Garcia-Fernández, J., Holland, L. Z., Williams, N. A. & Holland, N. D. *J. mar. biol. Ass. UK* **74**, 49–60 (1994).
6. Lacalli, T. C., Holland, N. D. & West, J. E. *Phil. Trans. R. Soc. B* **344**, 165–185 (1994).

Oldest rocks in Europe

Stephen J. G. Galer

How did the ancient continents differ from those of today, and how were they formed? These are key questions facing geologists studying the Archaean period, which spanned almost the first half of Earth history, ending 2.5 billion years (Gyr) ago. Answers can in principle be sought by examining old surviving rocks. But unfortunately, the further one looks back in geological time the more fragmentary the rock record becomes. It is for this reason that the discovery of yet older fragments of continental crust excites particular interest. On page 552 of this issue, Burton and colleagues¹ report convincing ages of 3.30 Gyr for two amphibolites — metamorphosed basalts — from the Lewisian complex of northwest Scotland.

From a worldwide perspective these ages, obtained by ^{147}Sm – ^{143}Nd and ^{207}Pb – ^{206}Pb dating, are not strikingly old: a number of Archaean terrains contain size-

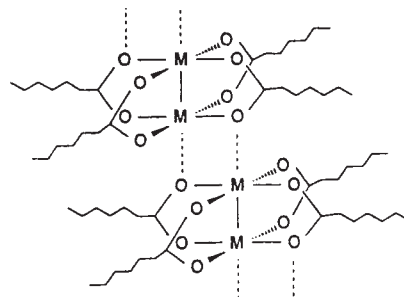
able pieces with ages in excess of 3.3 Gyr, notably in Canada, Greenland, southern Africa and Western Australia. The Acasta gneisses, located in the Slave province of northwestern Canada, contain components clocking in at 3.96 Gyr (ref. 2), making them the world's oldest rocks. Nevertheless, the results of Burton *et al.* are important for a number of reasons. The most trivial of these is that they have identified the oldest known European rocks, provided that Greenland is considered to be part of North America. Isolated pockets of 3.1–3.2-Gyr-old tonalitic gneiss are known from the eastern Baltic Shield^{3,4}, and previously held the European age record.

More important by far is that the history of crustal development in the Lewisian complex will have to be radically rethought. Geochronological investigations of the Lewisian have themselves had

Liquid crystal fans

THE birefringence of liquid crystals is familiar from digital watch displays, but their possibilities do not end there. Add a transition metal ion, either as a guest within an organic cage or as part of the molecular structure, and a host of magnetic, electronic and redox properties could come to light. So David V. Baxter and co-workers hope, at least. They

have been looking at compounds with the general formula $\text{M}_2(\text{O}_2\text{CR})_4$, where M is one of the group VI metals chromium, molybdenum and tungsten. The molecules are shaped like paddle wheels, and in the crystalline state form 'ladders' loosely held together by oxygen-to-metal intermolecular bonds as shown below (figures reprinted with permission from *J. Am. chem. Soc.* **116**, 4551–



4566; 1994; copyright American Chemical Society). Heat the crystals, and the different metal–oxygen bond strengths make themselves felt: Baxter *et al.* found that the weakly bonded tungsten compounds simply melted to an isotropic liquid, the strongly bonded chromium compounds formed a liquid crystal phase which decomposed on further heating, and many of the molybdenum compounds passed through all three phases in turn. The micrograph above, obtained on cooling isotropic liquid $\text{Mo}_2(\text{O}_2\text{C}(\text{CF}_2)_6\text{CF}_3)_4$, shows the fan-like optical texture of a hexagonal disordered discotic (stacked disk) liquid crystal. The authors conclude from optical, X-ray and NMR evidence that the stacking is retained here, but the intermolecular bonding is readily broken so that molecules can slip sideways or twist about their M–M axes. L. M.

a long history, following the classic study by Moorbath and colleagues⁵. In broad-brush terms, the consensus has been that the Lewisian crust was formed by differentiation of the mantle at about 2.9 Gyr and experienced high-grade granulite facies metamorphism around 2.7 Gyr ago (the grade refers to the severity of metamorphism, reflecting the peak temperatures and pressures reached). The new data from Burton *et al.* extend the geological history of the Lewisian back a further 400 Myr in time, at least in the Gruinard Bay section that they studied. Although Whitehouse⁶ reported a whole-rock ²⁰⁷Pb–²⁰⁶Pb isochron of 3.52 Gyr from the Lewisian of the Outer Hebrides, he dismissed this age, probably rightly, as being spurious. So why have previous studies failed to find rocks older than 2.9 Gyr? One answer might simply be that the 2.7-Gyr-old high-grade metamorphism at least partially reset the isotope clocks. The fact that Burton *et al.* find older ages in material metamorphosed under lower-grade conditions is suggestive of this.

Metamorphic resetting can only be part of the story, though, and the younger ages are clearly in most instances primary. Lewisian crustal evolution can now be traced from the Gruinard Bay amphibolites at 3.30 Gyr to the emplacement of the last members of the Scourie mafic dyke swarm 2.00 Gyr ago⁷ — a crustal stabilization interval comparable to that of the Kaapvaal craton in South Africa⁸.

Granitoid rocks of the trondhjemite-tonalite–granodiorite (TTG) suite make up much of the Lewisian complex. Burton *et al.* also report much younger ages of 2.40 Gyr for two trondhjemites from Gruinard Bay. But what are the genetic relationships, if any, between the amphibolites, trondhjemites and associated tonalites? Burton *et al.* view the amphibolites as surviving remnants of a once more extensive 3.3-Gyr-old crust of similar composition. Using their own and literature data, they are then able to make a strong case that both the trondhjemites and tonalites were derived by later intracrustal melting of such amphibolitic material, and that new mantle input was minimal. The argument is based on the initial ϵ_{Nd} of the respective isochrons (ϵ_{Nd} is the deviation in ¹⁴³Nd/¹⁴⁴Nd ratio at any time from the corresponding average ratio of the whole Earth, expressed in parts per 10⁴). Initial ϵ_{Nd} -values of the Lewisian rocks apparently decrease monotonically over time, starting from the amphibolite ϵ_{Nd} of +3 at 3.30 Gyr and passing to –4 for the trondhjemites at 2.40 Gyr. A secular change in ϵ_{Nd} like this is most easily explained by closed-system evolution of the crust at a ¹⁴⁷Sm/¹⁴⁴Nd ratio of about 0.13. As the depleted mantle is expected to have an ϵ_{Nd} of about +4 at 2.4 Gyr, juvenile mantle-derived material cannot have contributed significantly to the tonalites and trondhjemites.

The case for closed-system evolution is made far more compelling by the fact that the measured whole-rock ¹⁴⁷Sm/¹⁴⁴Nd ratios of the amphibolites themselves are around 0.13. A similar, albeit more complicated, story can be woven around the initial lead isotope ratios.

The TTG-suite rocks seem, then, to be genetically related to the amphibolites, presumably by melting. Support for this view comes from two other sources. First, it has long been supposed that the TTG-suite rocks in Archaean terrains require the presence of amphibole and garnet in their sources, on the grounds of their abundance patterns of rare earth elements^{9,10}. Second, experimental studies have shown that trondhjemitic and tonalitic magmas can indeed be produced by melting amphibolite^{11–13}.

Burton *et al.* have as yet identified only one isolated enclave of older mafic material in the Lewisian; with luck, more will be found in future studies. No doubt even older European rocks wait to be discovered as well. Nonetheless, a relatively complete and consistent picture of Lewisian geological history seems to be emerging: intracrustal melting of mafic precursors occurred over a protracted period and was extremely important in stabilizing the Lewisian crust and establishing its overall tonalitic composition. □

Stephen J. G. Galer is in the Abteilung Geochemie, Max-Planck-Institut für Chemie, Postfach 3060, 55020 Mainz, Germany.

1. Burton, K. W., Cohen, A. S., O'Nions, R. K. & O'Hara, M. J. *Nature* **370**, 552–555 (1994).
2. Bowring, S. A., Williams, I. S. & Compston, W. *Geology* **17**, 971–975 (1989).
3. Kröner, A. & Compston, W. *Contr. Miner. Petrol.* **104**, 348–352 (1990).
4. Lobach-Zhuchenko, S. B. *et al. Precamb. Res.* **62**, 375–397 (1993).
5. Moorbath, S., Welke, H. & Gale, N. H. *Earth planet. Sci. Lett.* **6**, 245–256 (1969).
6. Whitehouse, M. J. *Chem. Geol.* **86**, 1–20 (1990).
7. Heaman, L. M. & Tarney, J. *Nature* **340**, 705–708 (1989).
8. de Wit, M. J. *et al. Nature* **357**, 553–562 (1992).
9. Arth, J. G. & Hanson, G. N. *Contr. Miner. Petrol.* **37**, 161–174 (1972).
10. Martin, H. *Geology* **14**, 753–756 (1986).
11. Johnston, A. D. & Wyllie, P. J. *Contr. Miner. Petrol.* **100**, 35–46 (1988).
12. Rapp, R. P., Watson, E. B. & Miller, C. F. *Precamb. Res.* **51**, 1–25 (1991).
13. van der Laan, S. R. & Wyllie, P. J. *J. Geol.* **100**, 57–68 (1992).

ENZYME STRUCTURE

Cracking tyrosine phosphatases

John Tainer and Paul Russell

THE structure of a secret biological weapon responsible for the deaths of millions has been revealed. At the molecular level, this revelation should enable us to find out more about how phosphotyrosine signal transduction pathways are controlled. Of the two new protein tyrosine phosphatase (PTP) structures reported on pages 571 and 575 of this issue^{1,2}, one is a virulence factor for the bacterium of the genus *Yersinia* that is responsible for bubonic plague¹, infamous as the Black Death which killed one-third of Europe's population in the fourteenth century; the other is one of the first representatives from a class of low-molecular-weight PTPs ubiquitously expressed in mammalian tissues, but unrelated in sequence to the *Yersinia* PTP (ref. 2). These two new structures help solve the puzzle of how PTPs can have widely varying sequences and protein substrates, yet remain selective for phosphotyrosine and not phosphothreonine or phosphoserine.

These different PTP structures are striking examples of convergent evolution achieving highly similar active-site clefts. Both structures, as well as human PTP1B (ref. 3), which shares about 20 per cent sequence identity with the *Yersinia* PTP, have a common central structural core consisting of four parallel β -strands with surrounding α -helices. All three PTP

structures share a conserved phosphate-recognition loop (P-loop) structure which is formed by the signature sequence H/V-C-(X)₅-R-S/T-G/A/P (single-letter amino-acid code; X is any amino acid) common to more than 40 PTPs. Based upon crystal structures with bound tungstate¹ and sulphate², the P-loop provides a circular array of amino-nitrogen hydrogen bonds plus a critical arginine side-chain interaction which activate the catalytic cysteine thiolate and stabilize the subsequent bound phosphate anion. This P-loop lies at the carboxyl end of the parallel β -strands, forms the base of the catalytic cleft, and aligns the bound phosphate anion at the positive dipole of an α -helix. The surrounding loops are believed to confer the phosphotyrosine specificity of PTPs by producing a deep phosphate-binding site that prohibits tight binding by the shorter phosphoserine and phosphothreonine side chains.

Although the phosphate-binding P-loop structure is conserved, the β -sheet and α -helices in the protein core have distinct topologies. In the bovine low-molecular-weight PTP, the four β -stranded core region is formed by two β - α - β Rossmann motifs. In the *Yersinia* PTP structure, what would be the equivalent two β - α - β motifs are split by their β -strand interdigitation. Outside this core region, these structures vary greatly.

Binding site formation by a short contiguous sequence (the P-loop) and a structurally localized β -sheet core allow both sequence and flanking domain variations to provide exquisite control of substrate targeting and signal transduction while preserving the phosphotyrosine binding site. For example, the active-site face of the low-molecular-weight PTP contains aromatic residues as opposed to charged residues in the *Yersinia* PTP and the recently solved PTP1B structure³. In essence, these structures show how PTPs can make enormous distinctions in protein recognition and interactions while conserving their phosphotyrosine substrate and catalytic mechanism.

These PTP enzymes all use a closely related mechanism of catalysis: nucleophilic attack by a cysteine residue to form a thiol-phosphate intermediate, and phosphate displacement by a water molecule. The PTP1B crystals, soaked for two hours with tungstate and then frozen for data collection, showed concerted side-chain movements of P-loop cysteine and arginine side chains towards the tungstate³. The *Yersinia* PTP crystals cracked when soaked with tungstate, prompting Stuckey *et al.*¹ to crystallize the tungstate-bound enzyme; this allowed them to identify a major conformational change associated with the enzyme's reaction in a loop adjacent to the P-loop (see figure). Co-crystallization with tungstate shows that the loop containing aspartate at residue 356 folds over the active site and positions the aspartate side chain for proton transfer to the tyrosine leaving group. Thus in both the *Yersinia* and low-molecular-weight PTP structures, there is a second catalytically important loop that connects the end strand of the parallel β -sheet core to a helix and contains an aspartate positioned for proton transfer to the tyrosine-hydroxyl leaving group. As the low-molecular-weight PTP structure was solved with bound sulphate, it is not clear whether its aspartate position is pre-existing or inducible. Detailed characterization and analysis of this loop movement and the activity of this aspartate side chain in PTPs will undoubtedly be an important focus for further studies. Whereas this role for Asp 356 in proton transfer was predicted by site-directed mutagenesis⁴, the structure determination was necessary for the correct interpretation of the role of the active-channel Glu 290 in positioning the active-site arginine side chain.

Two other papers in press with *Biochemistry* establish X-ray⁵ and NMR solution⁶ structures of the bovine heart low-molecular-weight PTP and provide

important elaborations on the PTP structural analysis presented in this issue. These independent studies verify the existence of the P-loop structural motif and the central four-stranded β -sheet core. On the basis of chemical shifts observed in the presence and absence of phosphate and vanadate, the NMR structure⁶ confirms that phosphate binds to the P-loop at an α -helix amino terminus in solution. The X-ray structure at 2.2 Å resolution⁵ shows



View of the active-site conformational change that occurs upon ligand binding to the *Yersinia* protein tyrosine phosphatase. When the P-loop (gold) binds the phosphate analogue tungstate (red), a nearby loop (unliganded form, lavender; tungstate-bound form, white) moves about 7 Å and swings an invariant aspartic acid into the active site where it participates in phosphotyrosine hydrolysis. This change is also observed in the structure of the inactive Cys-to-Ser mutant enzyme complexed with sulphate (H. Schubert *et al.*, manuscript in preparation). Graphics by E. B. Fauman, H. L. Schubert and M. A. Saper (University of Michigan) using the programs MOLSCRIPT (P. Kraulis) and RASTER3D (D. Bacon, E. Merritt).

that the binding of phosphate, like sulphate in the structure of Su *et al.*², is aligned to the helix dipole, that the P-loop arginine side chain acts in recognition, orientation and stabilization of the bound phosphate, and that the aspartic acid on an adjacent loop is positioned for protonation of the tyrosine leaving group.

The structure determinations of these PTPs precedes identification of their *in vivo* substrates. The *Yersinia* enzyme possibly has several important targets, because the bacteria's strategy might be to disrupt multiple mechanisms that make up the host's defensive immune response. This could account for the broad substrate specificity and vigorous *in vitro* activity of the *Yersinia* enzyme. The biological functions of the low-molecular-weight PTPs are unknown, but this situation is likely to improve following the discovery of a fission yeast (*Schizosaccharomyces pombe*) gene (*stp1*⁺) encoding a low-molecular-weight PTP that is 42 per cent identical to its mammalian counterparts⁷. An important point is that the aspartate residue proposed to act as a proton donor is conserved in the yeast homologue. Fission yeast *stp1*⁺ was discovered because of its

ability to substitute for the tyrosine phosphatase activity of the cell-cycle regulator protein Cdc25, a dual-specificity phosphatase that dephosphorylates adjacent threonine and tyrosine residues in Cdc2, thereby activating Cdc2 kinase, which in turn promotes the G1/S and G2/M cell-cycle transitions⁸. The finding that low-molecular-weight PTPs are broadly conserved among eukaryotes provides reassuring evidence that these enzymes have important cellular activities that can now be explored using genetic approaches.

Dual-specificity protein phosphatases, including Cdc25 and Mkp1 (which dephosphorylates a MAP kinase⁹), make up a third group of PTP enzymes. These enzymes carry out biologically important dephosphorylations of both tyrosine and threonine residues, but their substrate specificity is extremely restricted. Apart from sharing a similar active-site loop sequence, dual-specificity protein phosphatases could hardly be more different from the tyrosine-restricted PTPs whose structures have been reported so far, hence determination of their structures is eagerly awaited.

Now that the molecular equivalent of the smoking gun for the Black Death has been unveiled, we can readily appreciate the difficulty of defending ourselves against this weapon. The common active-site structure for the PTP enzymes would make it difficult to distinguish friend from foe when designing drugs. In fact, the active-site core of the PTP enzymes apparently represents convergent evolution of a structural motif that stabilizes a cysteine thiolate, based upon the surprising structural similarity Stuckey *et al.*¹ have identified with rhodanese. These structures therefore mark a breakthrough in our understanding of two classes of PTP enzymes and establish the P-loop structural motif for the universal PTP signature sequence. □

John Tainer and Paul Russell are in the Molecular Biology Department, The Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, California 92037, USA.

1. Stuckey, J. A. *et al.* *Nature* **370**, 571–575 (1994).
2. Su, X.-D., Taddei, N., Stefani, M., Ramponi, G. & Nordlund, P. *Nature* **370**, 575–578 (1994).
3. Barford, D., Flint, A. J. & Tonks, N. K. *Science* **263**, 1397–1404 (1994).
4. Zhang, Z. Y., Wang, Y. & Dixon, J. E. *Proc. natn. Acad. Sci. U.S.A.* **91**, 1624–1627 (1994).
5. Zhang, M., Van Etten, R. L. & Stauffer, C. V. *Biochemistry* (in the press).
6. Logan, T. M. *et al.* *Biochemistry* (in the press).
7. Mondesert, O., Moreno, S. & Russell, P. *J. biol. Chem.* (in the press).
8. Miliar, J. B. A. & Russell, P. *Cell* **68**, 407–410 (1992).
9. Sun, H., Charles, C. H., Lau, L. F. & Tonks, N. K. *Cell* **75**, 487–492 (1993).

The hunt for Ras targets

Larry A. Feig and Brian Schaffhausen

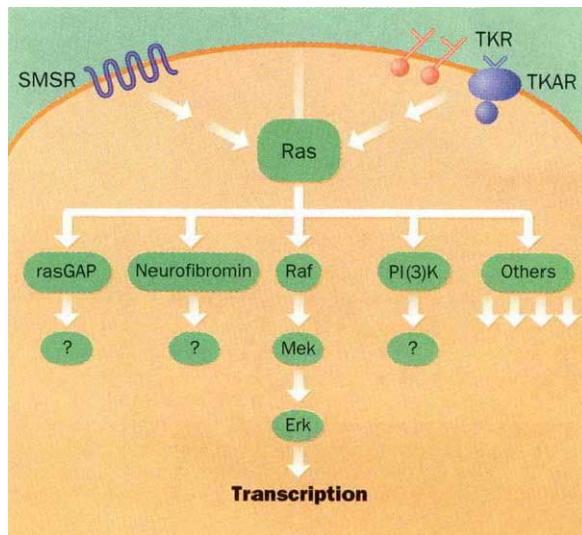
EVER since mutated *ras* genes were implicated in human carcinogenesis 12 years ago, it has been assumed that the proteins they encode regulate key signalling pathways involved in cell growth and development. Until recently, however, the hunt for direct downstream targets through which Ras acts has not been very satisfying. But now, remarkably, our cups are virtually running over with so-called 'effector' proteins that could transmit signals from Ras to various intracellular signalling cascades. The latest example, reported on page 527 of this issue¹ by Rodriguez-Viciana *et al.*, is phosphatidylinositol-3-OH kinase (PI(3)K), an enzyme already implicated in cellular growth control. This expands our view of Ras and its functions.

Ras proteins reside on the inner surface of the plasma membrane where they participate in transmitting signals from tyrosine kinase receptors and some receptors coupled to heterotrimeric G proteins². Ligand-bound receptors promote the active GTP-bound form of Ras, in most cases by enhancing the ability of guanine-nucleotide-exchange factors to accelerate the replacement of bound GDP with GTP. Ras proteins are then deactivated by interaction with GTPase-activating proteins (GAPs) that promote GTP hydrolysis by Ras. In some cell types, active Ras promotes cell proliferation; in others, it arrests cell division as it increases the expression of differentiated phenotypes.

Because they can bind active GTP-bound Ras in a region known to be required for downstream signalling, GAP proteins, including rasGAP and neurofibromin, were the first potential Ras effectors. A body of evidence supports the idea that rasGAP mediates some aspects of Ras function by complexing with other signalling proteins, although this pathway remains poorly defined³. A better understood pathway emanating from Ras involves the activation of an evolutionarily conserved cascade of kinases including Raf, Mek and Erk. The first step is binding of the Raf kinase to active GTP-bound Ras⁴, but it seems that Ras binding itself does not activate the intrinsic kinase activity of Raf: rather, it only localizes Raf to the plasma membrane where some unknown activation event takes place^{5,6}. For Raf then, Ras is a regulatable localization device, analogous

to GTP-binding translation factors.

Rodriguez-Viciana *et al.* now provide strong evidence that PI(3)K should be added to this growing list of Ras targets. They show that Ras binds to PI(3)K *in vitro* only when the Ras is in the active GTP-bound state. Moreover, mutations in the effector domain of Ras, or antibodies thought to block access to it, prevent interaction with the enzyme. Finally, the authors demonstrate that the



Downstream targets of Ras. Ras is activated in response to ligands that stimulate tyrosine kinase receptors (TKR), tyrosine kinase-associated receptors (TKAR), and some seven-membrane-spanning receptors (SMSR) that couple to heterotrimeric G proteins. It is becoming apparent that active Ras binds to a variety of proteins that may be responsible for carrying out its diverse functions. The best understood pathway emanating from Ras is mediated by the Raf kinase.

level of PI(3)K products generated *in vivo* can be increased or decreased depending on whether activated or dominant negative Ras mutants are expressed. Raf does not alter the level of these molecules, meaning that PI(3)K may be one component of a distinct parallel downstream pathway from Ras.

The PI(3)K studied by Rodriguez-Viciana *et al.* contains a catalytic subunit (p110) of relative molecular mass 110K, which binds Ras, and a regulatory p85 subunit. The enzyme phosphorylates phosphatidylinositols at the 3' position of inositol. The levels of the products, particularly phosphatidylinositol (3,4)-bisphosphate and phosphatidylinositol (3,4,5)-trisphosphate, have been tightly correlated with changes in the growth state of cells induced by either growth factors or oncogenes. Like Ras, however, these molecules seem also to be involved in signalling unrelated to growth, implying that they have a more diverse role in cell

regulation. This diversity may be derived in part from the multiple p110 and p85 subunits that have been detected^{7,8}. How many family members are regulated by Ras remains to be determined.

The cellular targets of the products of PI(3)K are not clearly defined, although protein kinase C family members may be included⁹. Attention has recently centred on the possible involvement of PI(3)K in membrane trafficking, because the p110 subunit of the enzyme shares substantial homology with VPS34, a protein involved in vacuolar sorting in yeast¹⁰. Along these lines, internalization of the receptor for platelet-derived growth factor (PDGF) seems to depend on PI(3)K activity¹¹.

Given the intense investigation of both PI(3)K and Ras, and the fact that a wide variety of cellular signals activate both proteins, one may wonder why this connection has not been observed before. One explanation is that two sets of experiments had already pointed to PI(3)K lying upstream — not downstream — of Ras in signalling pathways¹²⁻¹⁴. In one set^{12,13}, blocking PI(3)K recruitment to the plasma membrane in response to PDGF treatment of some cell types prevented Ras activation. These results implied that PI(3)K leads to Ras activation rather than the reverse. One interesting possibility that could resolve this apparent conflict is that PI(3)K competes with GAPs for binding to the Ras effector domain. In this way, it could actually increase the amount of Ras found in the GTP-bound state upon cell stimulation. Therefore, suppressing the ability of PI(3)K to be recruited to the plasma membrane where Ras resides

might increase the apparent activity of GAP and suppress Ras activation. This would place PI(3)K both upstream and downstream of Ras.

Another reason why Ras as an activator of PI(3)K has not been aggressively investigated until now may be that attention has concentrated instead on Rho proteins. These proteins are cousins of Ras and are involved in regulating the actin cytoskeleton. The p85 subunit of PI(3)K has a putative rhoGAP domain, suggesting that PI(3)K may be a target of Rho proteins. In fact, in platelets and fibroblasts, Rho function is required for PI(3)K activation in response to cell stimulation^{15,16}. One possibility is that Ras and Rho modulate the activity of different species of PI(3)K. Alternatively, Rho and Ras may cooperate to regulate PI(3)K activity. GAPs for both Ras and Rho proteins are physically associated upon cell stimulation³, so these two GTPases could exist together in a signal-

ling complex where they can coordinately regulate PI(3)K.

What then would Ras do in activating PI(3)K? At first glance, there seems to be no need for Ras to localize PI(3)K to the plasma membrane, as it does for Raf. The p85 subunit has two SH2 domains that bind to specific phosphorylated tyrosines on activated cell-surface receptors. In fact, the data of Rodriguez-Viciana *et al.* suggest that blocking Ras function does not inhibit membrane association of this form of the enzyme. However, a subtle role for Ras in localization may be involved. A non-transforming mutant of polyoma middle-T exists which lacks the ability to activate the SHC/GRB2 adaptors, suggesting that it fails to activate Ras in cells^{17,18}. Consistent with the new results of Rodriguez-Viciana *et al.*, the appropriate phosphoinositides are not produced¹⁹ *in vivo*. Interestingly, *in vitro* assays show that this mutant associates with the same amount of PI(3)K activity as wild-type middle-T. This suggests that Ras may function as a 'chaperone' to ensure lipids and enzyme interact in activated cells.

There may also be an SH2-independent signalling pathway involving Ras. A 110K subunit of PI(3)K that contains no 85K subunit has been detected in cell lysates²⁰. This form of the enzyme may need a membrane-localizing mechanism. By binding directly to p110 PI(3)K, Ras could provide this service.

Alternatively, Ras may not be responsible for getting PI(3)K to its substrates in the membrane. Rather, once the enzyme is there by virtue of its SH2 domain, the intrinsic activity of PI(3)K may be directly activated by Ras binding. This would be analogous to the ability of Ras to increase

the intrinsic activity of adenylyl cyclase in yeast. To the delight of those working on Ras, this would mean that mammalian Ras need not be merely a glorified membrane-localizing CAAX box after all! However, such an increase of enzyme activity upon Ras binding has not yet been demonstrated.

How many more targets of Ras are out there? When the two-hybrid system in yeast was used to screen a complementary DNA library for genes encoding Ras-

binding proteins, four different genes in addition to Raf were detected²¹. Although their importance has not yet been confirmed in mammalian cells, their existence suggests we have many courses to go before our feasting on Ras is complete. □

Larry A. Feig and Brian Schaffhausen are in the Department of Biochemistry and Sackler School of Graduate Biomedical Sciences, Tufts University School of Medicine, Boston, Massachusetts 02111, USA.

MUSICAL ACOUSTICS

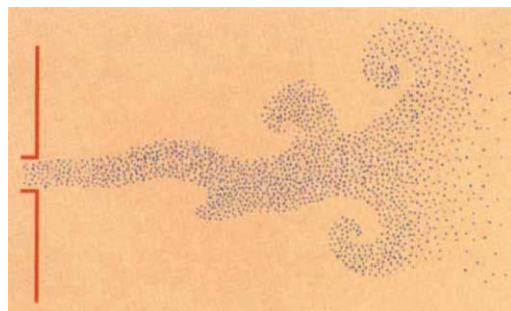
Flutes to hyperinstruments

Thomas D. Rossing

THE scientific study of musical instruments has attracted the attention of a number of famous scientists over the years: Rayleigh, Helmholtz, Tyndall, Raman, Savart and Saunders to name a few. But at no time in history have there been more exciting new developments in the study of musical instrument acoustics than we now see, due in part to the widespread availability of digital compu-

agates in the flow direction with a well-defined wave speed, and its amplitude grows in time, provided that its wavelength is not too short compared with the thickness of the jet.

Although this rather standard approach gives moderately good agreement with experiment in terms of blowing-pressure range, sound power output, acoustic spectrum, overblowing phenomena, initial transients and performance techniques, it is unsatisfactory in detail because of certain numerical deviations from experiment, and because of the ignorance concealed within the concept of a mixing region in which the jet flow becomes merged with the pipe flow; 95 per cent of the pneumatic energy carried by the jet into the pipe is dissipated in this mixing region (Fletcher).



Vortices developing in an air jet emerging from a slit.

The dynamics of the mixing region involve two different processes. In the first place, the jet

excites higher modes of the pipe with nearly the same efficiency as it does the basic plane-wave mode. As well as dissipating energy through viscous and thermal processes, these higher modes can also influence details of the jet interaction. The second process involves the interaction of vorticity, and particularly of concentrated vortices, with the acoustic flow (see figure). A vortex moving across an acoustic field can give energy to, or take energy away from, the field.

The physics of sound generation in reed woodwind instruments, such as the clarinet and saxophone, were also addressed in the woodwind session. Using a reactive power balance between the contributions of the various modes of the air column and those of the reed at the harmonics of the sound, it is possible to override the intrinsic nonlinearity of the flow-pressure relationship and to measure the reed stiffness when the instrument is actually played (Xavier Boutillon, Univ. Paris).

A numerical model of the air-reed

ters and other digital instrumentation for studying their complex behaviour. This fact was certainly brought home to a recent meeting of the Acoustical Society of America*.

A session on woodwind instruments covered the entire range from abstract theory to instrument manufacture and stimulated some interesting discussion between researchers, performers and instrument builders. Although we understand a great deal about the aerodynamic and acoustic interactions in flutes and other flute-like instruments, there are still gaps in our knowledge (Neville Fletcher, Australian National Univ., Canberra). The air jet produced by the player provides the prime excitation mechanism for all flute-like instruments, and does so because of a basic instability mechanism that was first discussed by Rayleigh. A sinusoidal disturbance on a plane jet prop-

- Rodriguez-Viciana, P. *et al.* *Nature* **370**, 527–532 (1994).
- Lowy, D. R. & Willumsen, B. M. *A. Rev. Biochem.* **62**, 851–891 (1993).
- Boguski, M. S. & McCormick, F. *Nature* **366**, 643–654 (1993).
- McCormick, F. *Curr. Opin. Genet. Dev.* **4**, 71–76 (1994).
- Leesters, S. J., Paterson, H. F. & Marshall, C. J. *Nature* **369**, 411–414 (1994).
- Stokoe, D. *et al.* *Science* **264**, 1463–1467 (1994).
- Hu, P., Mondino, A., Skolnick, E. Y. & Schlessinger, J. *Molec. cell. Biol.* **13**, 7677–7688 (1993).
- Parker, P. J. & Waterfield, M. D. *Cell Growth Diff.* **3**, 747–752 (1992).
- Liscovitch, M. & Cantley, L. C. *Cell* **77**, 329–334 (1994).
- Schu, P. V. *et al.* *Science* **260**, 88–91 (1993).
- Joly, M., Kazlauskas, A., Fay, F. S. & Convera, S. *Science* **263**, 684–687 (1994).
- Vallius, M. & Kazlauskas, A. *Cell* **73**, 321–334 (1993).
- Satoh, T., Fantl, W. J., Escobedo, J. A., Williams, L. T. & Kaziro, Y. *Molec. cell. Biol.* **13**, 3706–3713 (1993).
- Yamauchi, K., Holt, K. & Pessin, J. E. *J. Biol. Chem.* **268**, 14597 (1993).
- Zhang, J. *et al.* *J. Biol. Chem.* **268**, 22251–22254 (1993).
- Kumagai, N., Morii, N., Fujisawa, K., Menoto, Y. & Naramiya, S. *J. Biol. Chem.* **268**, 24535–24538 (1993).
- Dilworth, S. M. *et al.* *Nature* **367**, 87–90 (1994).
- Cambell, K. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **91**, 6344–6348 (1994).
- Ling, L. E., Druker, B. J., Cantley, L. C. & Roberts, T. M. *J. Virol.* **66**, 1702–1708 (1992).
- Shibasaki, F. Y., Homma, Y. & Takenawa, T. *J. Biol. Chem.* **266**, 8108–8114 (1991).
- Vojtek, A. B., Hollenberg, S. M. & Cooper, J. A. *Cell* **74**, 205–214 (1993).

*127th Meeting of the Acoustical Society of America, MIT, Cambridge, Massachusetts, 6–10 June 1994.

mechanism can be obtained by solving the Navier-Stokes equation for an incompressible fluid in the vicinity of the jet outlet (Peter Hoekje, Univ. Northern Iowa). The results of Hoekje's model confirm that mode locking and octave overblowing occur in an instrument whose frequency response has near-harmonic resonances, but not in one whose second resonance is much higher than twice the first resonance frequency.

"All well and good," responded Steven Wasser (Vern Q. Powell Flutes, Inc.), upon hearing these scientific papers, "but manufacturers of musical instruments are under constant pressure to improve productivity and enhance the performance qualities of their products." It is the responsibility of the manufacturer to establish a dialogue between the subjective artistic requirements of the musician and relatively objective manufacturing factors such as material properties, dimensions and processes. To illustrate his point, Wasser brought along Jerado di Stefalo, a professional flautist, to demonstrate the sounds of flutes made from gold and silver.



Ancient instrument, modern theory — the Flemish hammered dulcimer (P. Schenk, Collection Haags Gemeentemuseum).

Since 1986, the Hyperinstrument Group at the MIT Media Laboratory has been designing intelligent, interactive musical instruments that measure and interpret musical performance, among other things. Composer Tod Machover (MIT Media Center) discussed and demonstrated the current generation of hyperinstruments that incorporate MIDI controllers, samplers, custom signal processors, special input controllers and instrument sensors, and — of course — digital computers. Until 1991, Machover's hyperinstruments were designed primarily

for virtuosos such as cellist YoYo Ma, and emphasis was placed on measuring very subtle differences in interpretation and articulation from performer to performer. Since 1991, however, the hyperinstrument model has been extended for use by amateurs not proficient on an instrument, in an attempt to stimulate the imagination and allow room for growth and learning.

Mixing synthetic and acoustic sounds in a live performance requires the synthetic instruments to track the tempo and pitch of the acoustic performers in the same way an accompanist would. A computer model of human auditory sensing and pre-attentive processing can provide a good approximation of the conditioned stimuli with which the musical brain actually works (Barry Vercoe, MIT Media Center). A real-time model in which the first derivative of a constant- Q filter bank is fed into a temporal pattern detector to provide control for synthetic instruments needing to play in synchrony with the acoustic source was described. Live and videotaped demonstrations showed how well Vercoe's instruments can track variations in tempo and pitch by performing musicians.

Much can be learned about musical instruments by combining musical performance with scientific papers and acoustical discussions. The acoustical properties of hammered dulcimers, which originated in the Near East some 5,000 years ago and later became popular in Europe, China and the United States, have not been studied scientifically until very recently. A theoretical model of string vibrations in hammered dulcimers (illustrated here) has only just been developed (David Peterson, Central Arkansas Univ.), and the vibrational modes of the dulcimer soundboard and body have been studied by means of holographic interferometry and experimental modal testing (Gregory Canfield and T.D.R., Northern Illinois Univ.). The value of these models in understanding the dulcimer and its music were aptly illustrated by listening to the sounds of several hammered dulcimers, both by themselves and in ensemble with fiddle, guitar and voice, as in Southern folk music styles (David and Donna Peterson, Central Arkansas Univ., and Christopher Peterson, Washington Univ.).

Dialogue between performers, listeners and instrument makers is also important in understanding the acoustics of bowed string instruments. The Tokyo String Quartet offered a chance to compare the sounds of four different sets of instruments, their cherished old Italian instruments included. □

Thomas D. Rossing is in the Department of Physics, Northern Illinois University, DeKalb, Illinois 60115, USA.

DAEDALUS

Fractal money

MONETARY inflation, says Daedalus, is curiously popular. Steadily rising prices and incomes give a sense of progress. The endless pass-the-parcel rush as everybody tries to get rid of their money before it loses value is often mistaken for dynamism and economic vigour. An outmoded part of an inflating economy can decline discreetly without its turnover shrinking in numerical terms.

The trouble starts when the process accelerates into hyperinflation. The government prints more and more money, which loses value faster and faster. The traditional cure is a new currency. Sadly, this often catches the disease as well, so the process recurs.

Daedalus wants to tame this process by making it predictable. His model is the steady-state theory of the Universe. Galaxies expand apart; but as fast as they do so, new galaxies appear between them. The Universe inflates steadily for ever, always looking much the same. Similarly a currency (the pound, say) could depreciate steadily, predictably, until it was worth 10 per cent of its original value. The government would then decree a New Pound worth ten old pounds. It would exchange all cash and amend all accounts accordingly. This cycle would repeat endlessly.

As with all inflation, the government makes a steady profit. By printing money and spending it, it enriches itself while robbing its citizens. And a well designed steady-state inflationary regime, says Daedalus, could fund the government totally. Taxation could be abolished.

Modern taxation requires a vast state apparatus of questioning and snooping and grabbing. The oppressive, intrusive bureaucracy calls forth from the citizens even vaster amounts of form-filling and defensive accountancy. Freed from this massive unproductive burden, and with steady-state inflation safely stable and predictable, the real economy would surge forward as never before.

So Daedalus is wondering what steady-state inflation rate would allow the state to command half the citizens' wealth (a typical ambition these days). The naive answer is just 100 per cent, with money losing half its value per accounting period. Such an economy would replace its currency with new tenfold units every 3.3 such periods (with a small gain to the government each time: not all the old currency would be handed in). This answer ignores various subtleties; much hard-won experience will be needed to achieve the right steady-state inflation in practice. But the advantages for honest, fast-moving toil, and the disadvantages for cash-rich money-launderers, seem compelling.

David Jones

Foundations of gregariousness

SIR — Most benthic marine invertebrates have a pelagic larval phase, during which they may disperse widely. Once ready to metamorphose, the planktonic larvae of many species (for example mussels, oysters, barnacles, sand dollars) settle preferentially on or near conspecific adults, forming monospecific aggregations¹. Gregarious settlement has many advantages, most notably enhanced adult reproductive success², but at the cost of increased intraspecific competition. Aggregations must initially develop from a two-step process: solitary larvae first colonize an uninhabited substratum, then gregarious settlement occurs on or near these 'founders'. To date, research has focused on the latter step¹. It has been suggested that founders settle because they are unable to locate conspecifics and are incapable of prolonging their planktonic lives^{3,4}, a concept we term the 'desperate larva' hypothesis. One alternative, however, is that larvae differ in their substratum preferences and that founders are somehow distinct. We have found that females of a tube-dwelling polychaete worm produce larvae that settle in two different ways: one type colonizes uninhabited substrata (founders), whereas the other settles only in response to cues associated with conspecifics (aggregators).

We conducted still-water, single-substratum, laboratory settlement assays with larvae of *Hydroides dianthus*, a common member of epibenthic fouling communities along the east coast of North America. Three experiments were performed, each on a population of 90,000 larvae (pooled from the spawns of about 25 females) cultured in three 4-litre glass jars. We used two types of settlement substrata to assess larval responses: 'biofilmed' slides and 'adult' slides. Biofilmed slides were made by placing etched glass microscope slides into running unfiltered seawater for at least 5

days, so that they became coated with an organic/microbial film. Adult slides were treated similarly, but five small conspecific worms were attached by gluing their tubes to each slide. Two types of settlement assay were conducted until larval cultures were depleted: sample and whole-population assays. For sample assays, a sample of larvae was removed from culture jars and exposed to experimental substrata; for whole-population assays, all remaining larvae were exposed to substrata. Twelve replicate sample assays were run simultaneously in Petri dishes containing a single biofilmed or adult slide and 25 competent larvae. Assay dishes were examined for settlers after 24 h. For whole-population assays, the entire population of larvae was filtered down into a small volume and split evenly among nine replicate glass dishes, each containing a single biofilmed or adult slide. Dishes were placed on a shaker-table rotating at 50 r.p.m. for 1 h, and slides then examined for settlers. Larval types were defined by their responses: founders settled on biofilmed slides, and aggregators settled on adult slides.

In the first experiment, the whole population was exposed to both biofilmed and adult slides daily after day 3; therefore, all larvae had access to both substrata throughout development. Settlement on both experimental substrata began simultaneously as larvae became competent (Fig. 1), contrary to the prediction of the desperate larvae hypothesis. The bulk of settlement on both biofilmed and adult slides occurred soon after competence, with settlement occurring 7–8 days post-fertilization. Larvae continued to settle on adult slides in low numbers through day 26, but no larvae settled on biofilmed slides after day 14.

In a second experiment, the whole population was exposed to biofilmed slides daily (day 4 through day 14, weekly thereafter), then samples were exposed to adult and biofilmed slides. Under these constraints, only the larvae exposed to adult slides had the opportunity to settle gregariously, whereas the remainder of the population were forced to delay settlement, or settle on biofilmed slides. Surprisingly, virtually all founders settled between 5 and 12 days after fertilization, with the remaining larvae settling only in response to conspecifics (in sample assays) throughout a seven week period (Fig. 2). These results indicate the existence of a distinct subpopulation of larvae that responded to biofilmed substrata; once these individuals were removed, the remaining aggregators delayed settlement in the absence of acceptable conspecific-associated cues. In a similar experiment (not presented here) in which daily

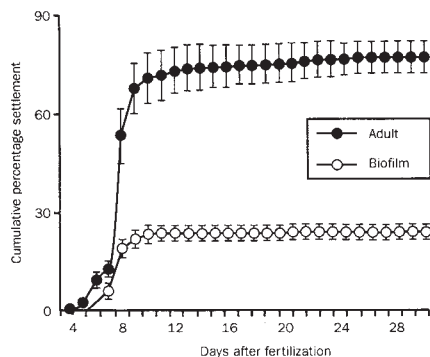


FIG. 1 Mean cumulative percentage settlement ($n = 9$, ± 1 s.e.) of *H. dianthus* when the whole population of larvae was exposed to biofilmed and adult slides daily.

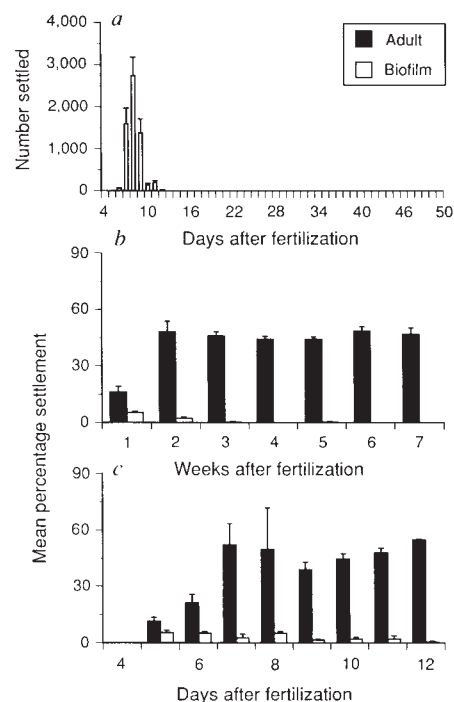


FIG. 2 Settlement of *H. dianthus* when the whole population was exposed to biofilmed slides daily, but only samples were exposed to adult slides. a, Number of founding settlers ($n = 3$, ± 1 s.e.) from daily whole-population assays of biofilmed slides. b, Weekly mean percentage settlement ($n = 12$, ± 1 s.e.) in sample assays of biofilmed and adult slides. Sample assays were conducted daily for the first 14 days, and weekly thereafter. c, Daily mean percentage settlement for the first week, a weekly mean of which is shown in b.

exposure of the whole larval population to biofilmed slides was omitted, sample assays revealed the same patterns as seen in Fig. 2 over a period of 70 days, after which, larvae began to die in culture rather than settle in the absence of conspecifics.

The desperate larva hypothesis was advanced on the basis of observations of the development of non-feeding pelagic larvae that exhaust a yolk reserve as they age in the plankton^{3,4}. If reduced substratum selectivity is a function of energetic desperation, rather than age alone, it should be possible to test this by starving the larvae of a planktotrophic (feeding) species, such as *H. dianthus*, and monitoring larval responses to substrata. We performed a third experiment, like the second, in which half the population of larvae was starved during days 10–14, and fed and starved subpopulations were

- Pawlik, J. R. *Oceanogr. mar. Biol. A. Rev.* **30**, 273–335 (1992).
- Leviton, D. R. *Am. Nat.* **141**, 517–536 (1993).
- Knight-Jones, E. W. *J. mar. biol. Ass. U.K.* **32**, 337–345 (1953).
- Wilson, D. P. *J. mar. biol. Ass. U.K.* **32**, 209–233 (1953).
- Pawlik, J. R. & Mense, D. J. In *Reproduction and Development of Marine Invertebrates* (eds Wilson Jr, W. H. Stricker, S. A. & Shinn, G. L.) (Johns Hopkins Univ. Press, Baltimore, 1994).

assayed separately. We found that settlement of founders could not be induced by starvation. Moreover, starvation resulted in a developmental reversion to a precompetent state, a condition that was reversed when larvae were again provided with food. An ontogenetic shift of this kind was recently described for the larvae of another species of marine polychaete⁵.

Besides random settlement, the desperate larva hypothesis is, to our knowledge, the only published explanation for the colonization of new habitats by larvae of gregarious species. The results of our study, a full version of which will be presented in a forthcoming paper, provide an alternative explanation for the process by which monospecific aggregations develop on hard substrata previously uninhabited by conspecifics.

Robert J. Toonen

Joseph R. Pawlik

Department of Biological Sciences and
Center for Marine Sciences Research,
University of North Carolina, Wilmington,
North Carolina 28403-3297, USA

Female platys prefer long tails

SIR — Female platys prefer a male with an artificial swordlike extension to his tail over the normal, swordless male of their own species¹. Swords are probably a primitive trait in the genus that has been secondarily lost by platys². How the preference for swords arose in the first place is not known, although it could be explained by virtually every extant theory of female-choice sexual selection³.

We explored the basis of female preference in platys, *Xiphophorus variatus*. We used life-sized, realistically coloured

wooden models. For each bout in a test series, a female was placed in the centre of a tank which was divided into three equal compartments by means of clear panes. Two mechanically yoked male models were used, one in each of the two end compartments. We scored the amount of time each female spent swimming against each of the clear dividers during two 10-minute periods. Most tests involved at least 20 individual female platys.

Females of this species of platy prefer model males with standard swordtails over models without such appendages. The preference, as measured by approach times, was 2.27:1 (see table); by way of comparison, a preference of 1.64:1 (measured by proximity times) has been recorded for another platy (*X. maculatus*)¹. Next we asked whether the preference was for the sword *per se*. We tested models with the sword placed at the top of the tail, extending forward from the head, lying along the midline of the body, or lying along the ventral surface of the body. In no case did the sword enhance the attractiveness of the model (see table).

To see if a less specific attribute of the sword is responsible for the effect, we tested a platy model with an exaggerated tail — that is, an elongated tail without a sword. Females did indeed prefer this model, and to the same degree as the swordtail model (2.28:1; see table); they showed no preference between an elongated-tail model and a swordtail model (1.04:1; see table).

We suggest that the general feature to which females are responding is tail length, as measured along the ventral surface. Thus the female preference observed here may simply be for males that are more obvious by virtue of the particular way in which the platy nervous system judges male size.

Female guppies — (*Poecilia reticulata*) close relatives of the platys — also prefer longer tails⁴. In nature male guppies tend to have short tails. Enlarged tails are not only more obvious to predators, they substantially slow male escape and probably raise the cost of locomotion. Perhaps the swords of swordtails represent a clever evolutionary compromise: enhanced attractiveness to females without the proportional increase in a male's visibility to predators and/or a concomitant reduction in his swimming speed or swimming efficiency. According to this interpretation, swords may themselves be a derived trait, and enlarged male tails may have been the primitive condition.

Sarah E Haines

James L. Gould

Department of Ecology
and Evolutionary Biology,
Princeton University,
Princeton,
New Jersey 08544,
USA

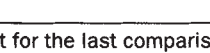
1. Basolo, A. L. *Science* **250**, 808–810 (1990).
2. Meyer, A., Morrissey, J. M., & Scharl, M. *Nature* **368**, 539–542 (1994).
3. Gould, J. L., & Gould, C. G., *Sexual Selection* (Freeman New York, 1989).
4. Bischoff, R. J., Gould, J. L. & Rubenstein, D. I. *Behav. Ecol. Sociobiol.* **17**, 253–255 (1985).

Earliest genets in Europe

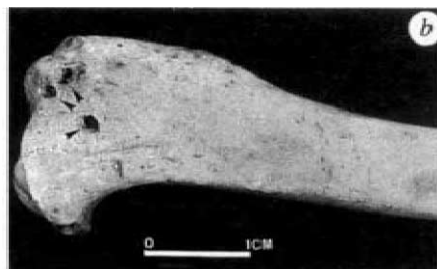
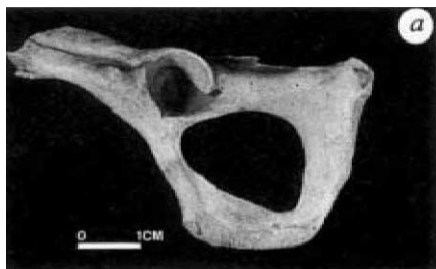
SIR — There is a theory maintaining that Charles Martel's founding of the short-lived Order of the Genet can be directly traced back to the stacks of genet pelts that his troops seized from the defeated Muslim armies after the battle of Poitiers in AD 732 (ref. 1). The dedication of an order to this animal is in itself peculiar, for the genet is a species of recent introduction to the European fauna, with no known fossil record, whose postglacial invasion of the subcontinent, whether by natural or anthropogenic means, is still open to debate^{2–4}.

A more interesting question relates to the putative domestication of the genet. Some authors mention genets as having been kept in southern Iberia and northern Morocco as a household-pest-destroying alternative to the cat^{5,6}, but not even the most comprehensive treatises on the subject dares to mention genets as potential domesticates, some even stating that this carnivore's solitary behaviour may not be suited to this manipulative process^{7,8}. Remains of different types of genets in African sites, both Maghrebi and sub-Saharan, have been systematically taken as evidence for hunting^{9,10}. In the context of this debate, the recent retrieval of a genet pelvis (*a* in the figure) from the Almohad levels of Mértola (Baixo Alen-

FEMALE PREFERENCES FOR MODELS

	normal male	Models compared	<i>n</i>	Mean time	<i>P</i>
		Ventral swordtail vs normal platy	22	185 81	0.013
		Dorsal swordtail vs normal platy	18	92 122	0.794
		Midline body sword vs normal platy	20	116 123	0.914
		Front-mounted sword vs normal platy	21	81 166	0.987
		Ventral body sword vs normal platy	20	176 120	0.185
		Elongated tail vs normal platy	17	281 123	0.022
		Elongated tail vs swordtail	18	143 158	0.381

Except for the last comparison, all models were tested against a standard male platy model (top left). Mean approach times are listed for reference. The *P*-values based on the signed-rank test (one-tailed: expectation that sword will increase attractiveness). *P*-values < 0.05 are printed in boldface.



a, Left hemipelvis of Genet (*Genetta genetta* L.), with the iliac wing fractured from the Almohad levels of Mintola (Baixo Alentejo, Portugal). b, Proximal portion of a left tibia from Iberian hare (*Lepus granatensis*, Rosenhauer) exhibiting the punctures (arrows) caused by the main cusps of genet's upper first molar (M1).

tejo, Portugal) constitutes a remarkable find. The bone was found lying among abundant microvertebrate remains filling a cesspit contextually dated to the first quarter of the thirteenth century AD and no later than 1238, the year of the conquest of the city by the Christians (S. Macias, personal communication). The sealed nature of this deposit precludes any possibility of the remains being intrusive. Prominent among the microvertebrate remains is the black rat (*Rattus rattus* L.), itself a subject of much archaeozoological debate and presumably one of the main reasons for the genet's putative domestication^{11,12}, but a more interesting find turned out to be a tibia from a hare (*Lepus granatensis* Rosenhauer) (b in the figure). This bone, which bears cut marks along its diaphysis indicating that it had previously been eaten by people, exhibits a series of chew marks on the proximal end, prominent among which are three on the medial side (arrows in part b) corresponding to the main cusps of a small carnivore's upper first molar (M¹). The protocone—metacone—paracone distances fully match the genet's M¹, rather than any mustelid of equivalent size and certainly have nothing to do with the M¹ from any cat. This all indicates that the viverrid probably got hold of the bone after it was discarded by people. (A.M., manuscript in preparation).

Although genets are not as anthropophobic as most European wild carnivores, readily venturing into poultry yards¹³, one would not expect to find

them in such a complex urban environment. We consider that this find has cultural connotations over and above the strictly palaeobiological.

As taphonomic and contextual evidence are circumstantial, we need unequivocal evidence to substantiate any claim of a domestication process having taken place at Mértola. The presence of the genet itself, on the other hand, offers physical proof in favour of the idea, long held by different scholars^{1,3,6}, that man introduced this species on Iberian soil and that this introduction was a quite recent zoogeographical event in the European subcontinent.

Arturo Morales

Laboratorio Arqueozoología,
Departamento Biología,
Universidad Autónoma de Madrid,
E-28049 Madrid, Spain

Climate and the pollen record

SIR — The matter of climate instability during the last interglacial period has been brought into focus by the GRIP ice-core results from the Summit, central Greenland, showing fluctuations between colder and warmer states¹. The validity of the last interglacial signal from the GRIP record has since been questioned by results from a duplicate core GRIP2 (refs 2, 3) and considerable further work may be necessary to evaluate the effects of ice flow on the stratigraphy of the two cores.

Broecker mentioned in passing in *News and Views*⁴ that further doubt has been cast on the issue by the absence of evidence for large climate shifts in the northern Atlantic marine or European pollen records. But this is not the case for the European pollen record.

A brief oscillation involving double peaks in *Abies* or *Picea* interrupted by expansions of *Pinus* before the onset of stadial conditions has been observed in European pollen records of the last interglacial period at Les Echets^{5,6}, Grande Pile⁷ and Devès plateau^{8,9} in France, and Sulzberg-Baden¹⁰ and Gondiswil¹¹ in Switzerland. The sensitivity of these sites

relative to others that do not seem to record this oscillation has been attributed to their proximity to ecotones⁶. A temporally more extensive vegetational reversion spanning approximately one-third of the interglacial succession is recorded in continuous pollen sequences from further south (Valle di Castiglione¹², Italy; Tenaghi Philippon^{13,14} and Ioannina¹⁵, Greece). Thus, two situations can be distinguished: (1) a moderate shift in vegetation after the thermophilous components have already disappeared towards the end of the last interglacial in certain sites of central Europe; and (2) a marked reversion after the middle of the interglacial when thermophilous trees spread again in southern Europe. The timing of these two types of changes cannot be independently established in the absence of precise radiometric dating, but their occurrence at different stages within the interglacial vegetation succession may point to the presence of two separate events.

Given the large physiographical, climatic and floristic variation between the different sites, we propose that such oscillations may be considerably more extensive than hitherto thought and should be viewed within the general framework of climate instability of the last interglacial. Thus, despite uncertainty over the climate mechanism(s) responsible for these oscillations and their chronostratigraphical relation to those defined within substage MIS-5e of the GRIP record, there is sufficient evidence to reject claims of absence of climate shifts in Eemian pollen records. Future high-resolution work on long continental sequences linking pollen analysis with other palaeoenvironmental studies will be able to provide further clues on the state of the last interglacial climate.

P. C. Tzedakis

K. D. Bennett

Sub-department of Quaternary Research,
Department of Plant Sciences,
University of Cambridge,
Cambridge CB2 3EA, UK

D. Magri

Dipartimento di Biologia Vegetale,
Università "La Sapienza",
00185 Roma, Italy

1. Hughes, A. *Revue Française Mammalogie* **2**, 52–60 (1928).
2. Kurtén B. *Pleistocene Mammals of Europe* (Aldine, Chicago, 1968).
3. Delibes, G. *Acta vertebrata* **IV** (1/2), 139–160 (1977).
4. Brosset, A. *Mammifères Sauvages de France* (Fernand Nathan, Paris, 1974).
5. Smit, C. J. van Wijngaarden, A. *Mammifères Menacés en Europe* (Conseil de l'Europe, Strasbourg, 1976).
6. Alcover, J. A. *Los Mamíferos de las Baleares* (Moll. Palma de Mallorca) (1979).
7. Mason, I. L. *Evolution of Domesticated Mammals* (Longman, London, 1984).
8. Clutton-Brock, J. *Domesticated Animals from Early Times* (Heinemann, London, 1981).
9. Romer, A. S. *Bull. Logan Mus.* **1**, 80–163.
10. Van Neer, W. *Ann. Sci. Zool.* **259** (Musée Royal de l'Afrique Centrale, Tervuren, 1989).
11. Erynck, A. thesis, Univ. Amsterdam (1989).
12. Armitage, P. West, B. & Steedman, K. *London Archaeologist* **414**, 375–383 (1984).
13. Delibes, M. *Acta vertebrata* **I** (1), 143–199 (1974).

1. Greenland Ice-core Project (GRIP) Members *Nature* **364**, 203–207 (1993).
2. Taylor, K. C. et al. *Nature* **366**, 549–552 (1993).
3. Grootes, P. M. et al. *Nature* **366**, 552–554 (1993).
4. Broecker, W. S. *Nature* **367**, 414–415 (1994).
5. de Beaulieu, J.-L., & Reille, M. *Boreas* **13**, 111 (1984).
6. de Beaulieu, J.-L., & Reille, M. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **72**, 147–159 (1989).
7. de Beaulieu, J.-L. & Reille, M. *Q. Sci. Rev.* **11**, 431–438 (1992).
8. Reille, M. & de Beaulieu, J.-L. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **80**, 35–48 (1990).
9. de Beaulieu, J.-L. & Reille, M. *Veget. Hist. Archaeobot.* **1**, 233–242 (1992).
10. Welten, M. *Eiszeitalter Gegenw.* **31**, 187–202 (1981).
11. Wegmüller, S. *Bull. Ass. fr. Etude Quat.* **1**, 29–34 (1986).
12. Follieri et al. *Pollen Spores* **30**, 329–356 (1988).
13. Wijnstra, T. A. *Acta bot. neerl.* **18**, 511–527 (1969).
14. Wijnstra, T. A. & Smit, A. *Acta bot. neerl.* **25**, 297 (1976).
15. Tzedakis, P. C. *Nature* **364**, 437–440 (1993).

Flight capabilities of *Archaeopteryx*

SIR — *Archaeopteryx lithographica* has a mixture of avian and reptilian features and so is not only an important link between two major vertebrate lineages but also offers a key to understanding the evolution of avian flight¹. Evaluations of muscle power² and osteology of the wrist³ both concluded that *Archaeopteryx* could not sustain powered (flapping) flight, but this is at odds with the asymmetry of its flight feathers, which is indicative of aerodynamic function^{4,5}. Here we compare feather vane asymmetry in *Archaeopteryx* with extant birds that fly, glide and are flightless.

Vane asymmetry varies with the position on the feather, the location of the feather in the wing, and the flight capability of the species. Before inferences of flight capability can be made, the other

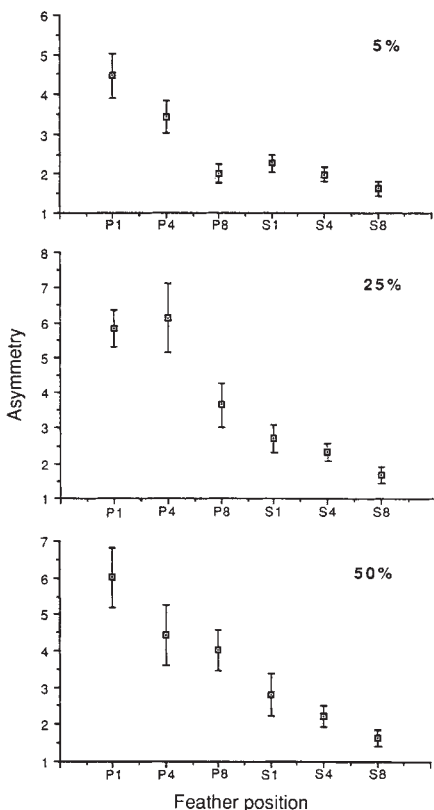


FIG. 1 Vane asymmetry (trailing-vane width divided by leading-vane width) in the wings of 17 species of flying birds drawn from 17 families, as a function of position on the feather (5, 25 and 50 per cent from tip) and feather position in the wing (P1, P4 and P8 represent the first, fourth and eighth primaries, counting the outermost feather as 1, and S1, S4 and S8 represent the first, fourth and eighth secondaries, respectively). Data shown are means plus and minus standard errors. Both feather position in the wing and position on the feather have significant effects on the extent of vane asymmetry. There is no significant interaction of these effects.

two sources of variation must be controlled. We examined the effects of feather position in wing, and position on the feather, on vane asymmetry in 17 species of flying birds (17 families). The first primary feathers were the most asymmetric (Fig. 1). At the point a quarter of the way up from the feather tip, the high ratio was maintained across the first four primaries; thereafter asymmetry was reduced. In 78 species (78 families) of flying birds (71 species using flapping flight, and 7 that predominantly glide) and 18 species (10 families) of flightless birds, we measured vane asymmetry 25% from the feather tip on the first or second (outermost) primaries (Fig. 2a–c). There was extensive asymmetry in the first primaries of both flapping and gliding birds, which was significantly greater than in the flightless birds (either including or excluding aquatic species).

Only three specimens of *A. lithographica* have the feathers sufficiently preserved to measure vane asymmetry: the Berlin and London specimens and the single feather. The single feather has an asymmetry of 2.2 at the 25% point, but where this feather comes from in the wing, and indeed whether it is from the same species, is unknown⁶. In the Berlin and London specimens the feathers are embedded in wings and therefore overlap each other. The distal ends of some of the feathers are sufficiently exposed to make measurements of vane asymmetry at the 25% position from the feather tips. We measured primary feathers 4, 5 and 6 of the left wing of the Berlin specimen and primary feathers 3 and 4 from the left wing of the London specimen. Vane asymmetry averaged 1.44 for the London specimen and 1.46 for the Berlin specimen (Fig. 2d). The asymmetry in the primary four feathers of *Archaeopteryx* is not significantly different to that of modern flightless birds (t -test, $t=0.94$, $P=0.36$) and substantially and significantly lower than that of modern flying birds (whether flapping ($t=12.52$, $P<0.0001$) or gliding ($t=5.26$, $P<0.001$)). The asymmetry is considerably lower than the lowest value (2.22) at equivalent positions for any flying bird across the 78 different families considered.

The extent of asymmetry in the feathers of *Archaeopteryx* is only slight, signifi-

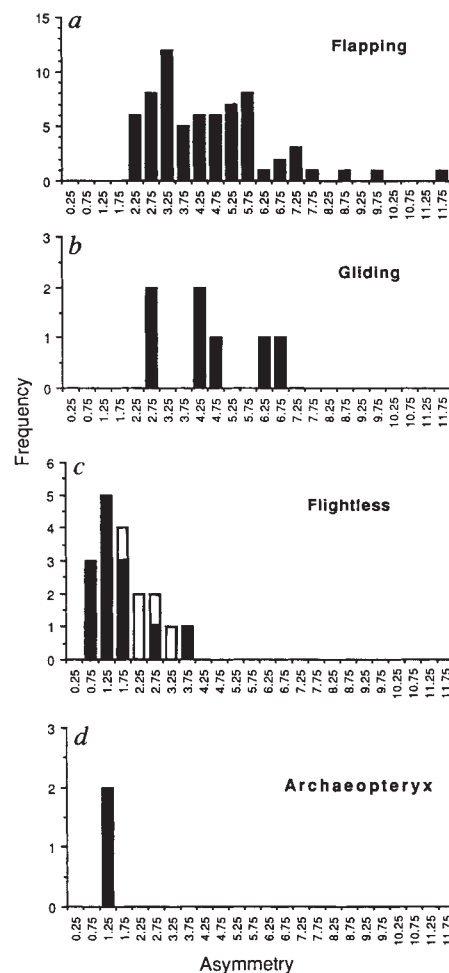


FIG. 2 Frequency distributions of vane asymmetry (trailing-vane width divided by leading-vane width) 25 per cent from the tip of the feather for the first or second primary of *a*, 71 bird species that use predominantly flapping flight; *b*, 7 using predominantly gliding flight; and *c*, for 18 species of flightless birds. Data for flightless aquatic birds are shaded; non-aquatic flightless species are unshaded. *d*, Vane asymmetry 25 per cent from the feather tip for the fourth primary of the Berlin and London specimens of *Archaeopteryx*.

cantly lower than in modern flying birds and comparable to that of extant non-flying birds. Our feather asymmetry results are thus consistent with data suggesting that its muscles were insufficient to power sustained flapping flight² and that the structure of its wrist could not execute a flapping motion in the wings³.

J. R. Speakman

S. C. Thomson

Department of Zoology,
University of Aberdeen,
Aberdeen AB9 2TN, UK

Scientific Correspondence

Scientific Correspondence is a relatively informal section of *Nature* in which matters of general scientific interest, not necessarily those arising from papers appearing in *Nature*, are published. Because there is space to print only a small proportion of the letters received, priority is usually given according to general interest and topicality and to contributions of fewer than 500 words.

1. Olson, S. L. & Feduccia, A. *Nature* **278**, 247–248 (1979).
2. Speakman, J. R. *Evolution* **47**, 336–340 (1993).
3. Vazquez, R. J. *J. Morphol.* **211**, 259–268 (1992).
4. Feduccia, A. & Tordoff, M. *Science* **203**, 1021–1022 (1979).
5. Norberg, R. A. in *The Beginnings of Birds* (eds Hecht, M. K.) 303–318 (1985). Freunde des Jura-Museums Eichstatt.
6. Owen, R. *Proc. zool. Soc. Lond.* **12**, 272–273 (1864).

The physicist's tale

Walter Gratzer

La vie à fil tendu. By Georges Charpak with Dominique Saudinos. *Éditions Odile Jacob*, Paris: 1993. Pp. 240. FF130 (pbk).

YOUTH, wrote Benjamin Disraeli, is a blunder, manhood a struggle and old age a regret. Yet not so for Georges (formerly Grisha) Charpak: youth for him was a fearsome test of moral fibre and endurance, maturity, it seems, a serene progress from conquest to conquest, and the advancing years have brought only a meagre satisfaction.

Scientists' lives (as has been abundantly demonstrated in recent years) are only rarely worth the telling. They capture the reader's attention when the world beyond the confines of the laboratory intrudes or when an anecdote reveals something unexpected, preferably discreditable, about some illustrious contemporary. Charpak's early life was indeed interesting, for, like other Central European Jews of his generation, he was caught up in the maelstrom of historical events and, unlike most others, survived. He grew up, a Yiddish speaker, in a *shtetl* in the Polish Ukraine between the wars. His parents sought escape from poverty and an uncertain future by emigrating, and after some false starts, including a period in Palestine, landed in Paris.

Charpak discovered early on a taste for mathematics and got a precarious foot on the academic ladder simply by presenting himself at the door of the nearest *lycée* and demanding admission. The privations of life in a grindingly poor immigrant community were made tolerable for the young Georges by his membership of a communist youth organization, the Red Falcons, and then — when the Soviet–German pact brought, as to so many other idealists, disillusion with Stalinist rhetoric — the more wholesome *Auberges de Jeunesse*. There he found around the camp-fire a companionship and solidarity that sustained him in the trials to come. For Charpak, stung by the antisemitism of many of his school-fellows under the German occupation, and eager for the good fight, joined a resistance group, was caught by the Vichy police, imprisoned and not long after sent to Dachau.

Tall, blue-eyed and now Charpentier, he passed easily enough for a *goy*, except for one hideous moment when he went to sleep on his feet and was heard mouthing Yiddish by a passing German officer. His hot denials were accepted and the discipline and selflessness that distinguished him and his group of friends saw him through. He gives a moving account of what he observed during this gruesome time.

The invasion of Hungary in 1956 imbued Charpak with a terminal disgust for communism, and dismay at the venality and incompetence of the Fourth Republic made him resolve henceforth to concentrate his energies on physics. French physics at that time was largely stuck in the

sics) and it was there that he made the majority of his most celebrated discoveries, which led him eventually to the handshake with the King of Sweden (to universal acclaim, as I understand, from the particle-physics community). After Stockholm there follow accounts of the usual pleasures of the Nobel laureate's pampered life — summer schools and lectures in agreeable places, eye-opening encounters with exotic peoples, good works for deserving causes and even a little more science.

A contemporary historian once remarked that the advantage of autobiography is the opportunity it gives of telling

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Georges Charpak at CERN in 1992.

nineteenth century. Charpak is discreet in his account. Theoretical physics was dominated by the duc de Broglie, who ran a group absorbed in abstractions with little bearing on the stirring events that were unfolding in physics laboratories elsewhere. A gulf separated theoretical and experimental physicists. Jeremy Bernstein in *The Life it Brings* has portrayed the place and time with greater candour and aplomb. Louis Leprince-Ringuet, who, Charpak says, directed a world centre of excellence in the study of cosmic rays, was according to Bernstein one of the last of the Baconians, untroubled by the constraints of theory: you had only to put your photographic plate in the right place, and, in Bernstein's words, God would provide.

Charpak finally found a patron in Frédéric Joliot, by then much preoccupied with scientific politics, and he began to develop a passion, which has endured throughout his career, for the design of particle detectors, with at their heart the "taut wire" of the title. Charpak's scientific life was mostly centred on CERN (European Laboratory for Particle Phy-

the truth about other people. Charpak does not see things in this way: he is courteous about everybody in physics, even the dead, and he gives no hint of the mephitic atmosphere that, according to other accounts (Gary Taubes's *Nobel Dreams* for example) pervaded much of CERN. There are some decent anecdotes — the radioactive overcoat, disposed of in gobbets in dustbins around the city, the human head, preserved in formaldehyde and carried around in a sports bag, for research on the development of an imager, and a few more.

We are left then with a self-portrait of a clearly exceptional scientist and a genial man, who has lived an admirable life. He is also one of the last of a remarkable breed; a generation born of adversity and hungry for intellectual nourishment, has faded away. A new and equally determined brood has emerged from quite other lands, and to them the torch is now passing. □

Walter Gratzer is in the MRC Muscle and Cell Motility Unit, Kings College, 26–29 Drury Lane, London WC2B 5RL, UK.

CERN Photo

Green networks

Timothy O'Riordan

Fantasy, the Bomb and the Greening of Britain. By Meredith Veldman. Cambridge University Press: 1994. Pp. 325. £35, \$54.95 (hbk); £12.95, \$17.95 (pbk).

AT first sight there does not seem much to connect the fantasy writings of C. S. Lewis and J. R. Tolkien with the Campaign for Nuclear Disarmament (CND) and the rise of environmentalist protest in Britain. It all sounds like a quiz for the radio programme 'Brain of Britain'. But Meredith Veldman creates a most plausible intellec-

of control. British imperial power has waned, but its citizens still yearn for the time when Britain was a force in the world. So the fantasy is complicated by a paradoxical interpretation of social change, where adherents of the new order search for some areas of speciality where Britain can once again lead the world. Artistry, intellectual inventiveness and caring for humanity on a peaceable Earth could be such causes. This mood does intertwine the three themes of Veldman's book. The search for a better coexistence is a powerful connector.

The third idea is the desire to make this new world compatible with human values. This means getting on top of big technology, big government, big industry and big science. This is where Veldman's analysis

suffers from what seems like an artificial conjunction. The environmental movement is far more complicated than these simplistic 'antis'. So too are the convoluted writings of Tolkien and Lewis. The CND movement was genuinely fearful of true holocaust, but it was not especially anti-bigness so long as it could be harnessed in the cause of peace and goodwill.

This is a clever book, well written, thoughtfully presented and subtly connected. In each of the three areas of primary analysis there is work of authority and scholarship. One cannot but be impressed by the brave attempt to connect fantasy, anti-nuclearism and the early Green movement into a single strand. But one could connect many other areas of the arts, political protest and social reform if one looked hard enough. The author's synthesis is not in itself a fully convincing argument, but it is certainly a cleverly presented one.

What makes the environmental movement tick is not

its romanticism but its hard-headedness. It is a thoroughly sophisticated operation that has always been grounded in a mixture of science, politics and an appeal to inner morality.

One senses the same of CND. Science has always played a powerful role in these movements, but always through a side-ways look. This is surely their strength and their legacy. □

Timothy O'Riordan is in the School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, UK.

New in paperback

Art Restoration: The Culture, the Business and the Scandal by James Beck with Michael Daley (John Murray, £12.99). A controversial attack on the conservation profession, holding it to be at the mercy of scientists and attribution-seeking curators. The argument failed to convince *Nature's* reviewer, who said that it was "seriously undermined by selective and misleading use of evidence" (Ian McClure, *Nature* **366**, 521; 1993).

Flatland by Edwin A. Abbott (Oneworld, £4.95, \$7.95), a reprinting of this perennial science-fiction classic, first published more than a century ago.

Animal Minds by Donald R. Griffin (University of Chicago Press, \$14.95, £10.25). The book "has no equal in the sheer range of species, studies and behavioural phenomena that [Griffin] draws together" (Andrew Whiten, *Nature* **360**, 118; 1992).

Kepler's Physical Astronomy by Bruce Stephenson (Princeton University Press, \$14.95), an analysis of the development of Kepler's laws.

Vital Signs 1994–1995: The Trends That Are Shaping Our Future by Lester R. Brown, Hal Kane and David Malin Roodman (Earthscan, £10.95). The third annual edition of this popular source of environmental information from the Worldwatch Institute.

Sir Joseph Banks: A Global Perspective edited by R. E. R. Banks, B. Elliott, J. G. Hawkes, D. King-Hele and G. I. Lucas (Royal Botanical Gardens, Kew, £12). A collection of 20 papers arising from a conference held last year at the Royal Society, London, to commemorate the two-hundred-and-fiftieth anniversary of the birth of the British botanist.

Evolution in Age-Structured Populations by Brian Charlesworth (2nd edn) (Cambridge University Press, £19.95, \$29.95). Of the first edition, published 15 years ago, *Nature's* reviewer wrote: "An important book which can be recommended to anyone with an interest in population genetics and evolutionary biology".

Dynamics of the Standard Model by John F. Donoghue, Eugene Golowich and Barry R. Holstein (Cambridge University Press, £25, \$39.95).

Quantum Field Theory by Lowell S. Brown (Cambridge University Press, £25, \$39.95).

Chemical Oscillations and Instabilities: Nonlinear Chemical Kinetics by Peter Gray and Stephen K. Scott (Oxford University Press, £25). Intended for chemists, mathematicians and engineers.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

London protest — the Campaign for Nuclear Disarmament has always been grounded in a mixture of science, politics and an appeal to inner morality.

tual case, even though its practical application may prove less persuasive.

Veldman seeks to make the connection through three interrelated schemes. The first is some sort of arcadian myth that the old ways are being split asunder by technology, big power politics and a general lack of social care in government, industry and science. This activates a romantic antisecularism that was always embedded in British culture. Various events trigger it, and nuclear power in all its manifestations is one of them.

The second theme is the sense of a loss

The Hermetic Hamlet

Ryan J. Huxtable

The Jewish Alchemists: A History and Source Book. By Raphael Patai. Princeton University Press: 1994. Pp. 617. \$35, £29.95.

CLASSICAL knowledge was preserved in the expanding Arab world during the dark ages in western Europe following the resorption of the Roman empire. Much of this knowledge was rediscovered and reacquired by the West in the centuries thereafter. This passage is memorialized linguistically by the number of words still used that are prefixed with the Arabic definite article, *al*: alcohol, alembic, algebra, alchemy, algorithm, almanac, alizarin, alkali. None conjures up the mystical origins of Western thought so well as the word alchemy. The word derives from the Greek word for Egypt, the presumed home of the suppositious Hermes Trismegistus. He was the factitious author of the body of writings from which hermetic philosophy derived, and which in turn fuelled the neo-Platonic revival led by Marsilio Ficino and Pico della Mirandola in the fifteenth century. Hermeticism was closely allied to alchemy, with its mystical beliefs in the alkahest, or universal solvent, the panacea, or universal cure, and the philosopher's stone, which converted all it touched to gold.

These recondite beliefs reflect the moral basis of alchemy. It is built on the theory that all forms of matter are variants of the one basic essential substance. As Shakespeare phrases it in one of his many Hermetic touches: "A king may go on a progress through the guts of a beggar" (*Hamlet*, Act IV). Hence base metals can be transmuted to precious metals (or vice versa), and a malfunctioning body could be healed by an 'essence' of a plant. This is the Platonic idea, still so important in Western thought, that all manifest forms are but imperfect aspects of a pure, idealized form — hence the word 'idea', which has been debased over the centuries to something very different from what Plato intended. In alchemy, the distinction between the animate and inanimate was blurred (anima, of course, being the soul, or spirit). If man consisted of liquid, solids and spirit, so did copper, only in different proportions.

Alchemy as a moral system explains the otherwise puzzling problem of how a belief in the philosopher's stone could survive centuries of failed demonstrations, and how rulers and the power élite could continue to employ alchemists: the success of an alchemist depended not so

much on his last experiment as on his moral status. Indeed, alchemical doctrine equated the philosopher's stone with Christ himself and Hermes Trismegistus ("the three-times greatest") with the Christian trinity. Belief in the one inherently involved belief in the other.

The Hellenic and Christian roots of alchemy are obvious. Raphael Patai, in this volume dense with exegesis and translation of original documents from Hebrew and Arabic, argues also for its Jewish roots. There was a mediaeval belief that Moses was an alchemist, and one of the important early alchemists, although perhaps as apocryphal as Hermes Trismegistus, was Maria the Jewess. She is hailed as the founder of both alchemy and the Hermetic art. Maria is supposed to have invented the water bath for heating substances that is still known as the *bain marie* in France and the *Marienbad* in German. The writings ascribed to her also contain the oldest description of a still. Like Hermes Trismegistus, both Moses and Maria came from Egypt, the *Khemia* of

the Greeks and the *al-kimia* of the Arabs. Indeed, there was even a belief that Moses and Hermes Trismegistus were one and the same person. Patai moves slowly forward from the times of mists and legends to the nineteenth century. Shadowy figures, who, if they ever existed, we know about only through the writings of others, make their vague appearances and vanish. Obscure but earnest documents are translated that discuss at length the unsubstantiated beliefs of a pseudoscience largely ignored by the modern world. The biographies of Jewish practitioners of alchemy are scrutinized through the centuries. We learn that the great physician Avicenna, who showed that diabetic urine was sweet, was supposedly trained by Jacob the Jew. Occasionally a brightly lit figure steps from the shadow. Raymond de Tarrega is one. He was murdered in his prison cell in Spain for heresy in 1371. The plants used by de Tarrega are still familiar to any ethnopharmacologist: opium, camphor, hemlock, ebony, mullein. De Tarrega seems to have been a fourteenth-



W. D. MATTHEW (left), curator in the American Museum of Natural History's Department of Vertebrate Paleontology, circa 1916. Taken from *The American Museum of Natural History's Book of Dinosaurs and Other Ancient Creatures* by Joseph Wallace (Simon and Schuster/Prion, \$25, £14.99), a popular well-illustrated account of some of the museum's most important fossil vertebrate specimens and their historical background. The book is published to accompany the major renovation of the museum's fossil vertebrate galleries, due to be completed in 1996.

century precursor of Giordano Bruno. The latter hermeticist was burned in Rome on the Campo di Fiori in 1600 for his Copernican and Hermetic beliefs, thereby giving Galileo a powerful impetus to recant when he was similarly accused a scant 32 years later. The Bible and the Talmud are painstakingly examined for alchemical references.

Finally, Patai asks whether there were any differences between the alchemy practised by Gentiles on the one hand and Jews on the other. In other words, was there such a thing as Jewish alchemy? The book opposes the general anti-alchemical bias of Jewish historians, who, the author argues, play down the involvement in the art of their co-religionists.

This is a work of scholarship. In reading it, however, I cannot suppress a niggling thought that arcane matters may occasionally be best left arcane. One emerges from the chore with a sensation of having dined on tofu with Miss Havisham and a head full of dusty words such as molybdochalkon, aurum potabile, orpiment and magisterium. □

Ryan J. Huxtable is in the Department of Pharmacology, College of Medicine, University of Arizona, Tucson, Arizona 85724, USA.

■ Princeton University Press has also just published *The Business of Alchemy: Science and Culture in the Holy Roman Empire* by Pamela H. Smith. \$45, £30.

Gene hunting in the raw

Peter Goodfellow

The Gene Hunters: Adventures in the Human Genome Jungle. By William Cookson. Aurum: 1994. Pp. 208. £16.95.

WRITING a book about science for a general audience is not easy. The author has to translate strange language and concepts into everyday words and metaphors. Moreover, the book should be entertaining. Frequently the task is regarded by publishers as so difficult that it can be undertaken only by a journalist. If the journalist does not have a few friends willing to correct the scientific errors, the results can be amusing. Publishers concerned with accuracy are usually forced to employ scientists to write about science. The drawback with this approach is that most of us who would lay claim to the title of scientist are so frightened about the opinions of our colleagues that comprehensibility is readily sacrificed for accuracy. Obviously there are exceptions to this general rule and occasionally scientists have presented science in a human and entertaining way.

The Gene Hunters is a 'kiss and tell' book from a scientist involved in gene mapping and positional cloning experi-

ments. According to the book's cover: "Dr Cookson takes us inside the world of the gene hunters to show us science in the raw." The heart of the book is a description of Cookson's attempts to identify genes involved in asthma. These chapters have a disarming ring of truth. The author admits his fallibility and bares his sins and successes. He also points his finger at his enemies — and stabs them in an orifice with it. If this book was a newspaper or a television programme, several people would be claiming the right to reply. The other chapters are much less convincing. The early chapters set the stage with descriptions of DNA, genetics and gene mapping. Next come anecdotal descriptions of cloning of genes causing cystic fibrosis, Duchenne muscular dystrophy and maleness. These chapters convey a view from the balcony rather than centre stage. Cookson is always opinionated, which makes the book a good read, but he is often wrong. He is entitled to the opinion that everyone has forgotten the name Bateson although bigots from Cambridge may take a different view. He is not entitled to commit scientific errors to print. Chromosome walking in 1985 was not restricted to plasmids and, in any case, plasmids can be used to clone fragments much bigger than 2,000 base pairs. The mammalian sex-determining gene is not 240 base pairs long; the open reading frame alone is three times that size. Peter Goodfellow cannot be described as having a puckish sense of humour as it is well known that he is humourless.

The book finishes with a brief description of the genetics of cancer and a consideration of the ethics of molecular genetics and the looming problem of the new eugenics. Again, the voice that speaks has a unique flavour. It is only a slight exaggeration to suggest that Cookson blames Galton, the British class system and men with neat beards and fastidious dress sense, in that order, for the evils of eugenics. After roundly and correctly castigating eugenics, he suggests that the "more reliable" way to choose genes for intelligence for your offspring is to "find a potential mate with the abilities you seek and to persuade them to father or mother your children". Unless I am mistaken, this is a flawed strategy based in false eugenic principles.

I congratulate the author on having tried to express his view of science in a language that is lively and easy to read. I am disappointed that he did not take more care to check his scientific facts. Overall the book is entertaining and can be read in a deck chair on a beach. But it fails the final test of a popular science book: I would not give it to my mother. □

Peter Goodfellow is in the Department of Genetics, University of Cambridge, Downing Street, Cambridge CB2 3EH, UK.

Feynman and the art of physics



"WHEN my calculations didn't work out, I would watch the girls... I'd like to use the place to work in," Richard Feynman said of a topless bar he used to frequent. He was testifying in court on behalf of the bar's proprietor, who had been charged with permitting "lewd performances". As this page from one of his notebooks shows, he also practised his sketching there. The picture appears in *No Ordinary Genius: The Illustrated Richard Feynman* (Norton/Weidenfeld and Nicolson, \$29.95, £18.99) by Christopher Sykes, a collection of personal reminiscences by Feynman's family and friends.

The role of magmas in the formation of hydrothermal ore deposits

Jeffrey W. Hedenquist & Jacob B. Lowenstern

Magmatic fluids, both vapour and hypersaline liquid, are a primary source of many components in hydrothermal ore deposits formed in volcanic arcs. These components, including metals and their ligands, become concentrated in magmas in various ways from various sources, including subducted oceanic crust. Leaching of rocks also contributes components to the hydrothermal fluid—a process enhanced where acid magmatic vapours are absorbed by deeply circulating meteoric waters. Advances in understanding the hydrothermal systems that formed these ore deposits have come from the study of their active equivalents, represented at the surface by hot springs and volcanic fumaroles.

WHAT are the sources of metals concentrated in ore deposits? This question was first addressed in the sixteenth century by Agricola, who concluded that heated rainwater leaches metals from rocks and then transports the metals to the sites of ore deposition. A century later, Descartes proposed that vapours released during cooling and crystallization of the Earth's interior are responsible for filling fractures with ore. Thus, the stage was set for controversy between the Neptunists and the Plutonists, a debate that continues today.

Hydrothermal systems powered by magmatic intrusions dominate fluid movement in the upper crust¹, and are responsible for convecting a large proportion of Earth's heat to the surface. At the same time, the fluids transport metals, forming the single most important class of ore deposits, those of the hydrothermal category. For example, large, low-grade deposits of copper (plus molybdenum or gold) within shallow igneous intrusions, known as porphyries, provide more than half of the world supply of copper and molybdenum. Other types of hydrothermal ore deposit (Table 1), some of which form at a greater distance from the magmatic hearth, are also significant repositories of gold, silver, lead, zinc, tin, tungsten and other elements. Non-magmatic hydrothermal systems are also important in metallogenesis, such as the sedimentary-hosted deposits rich in copper, lead and zinc that form by compactive expulsion of fluids from large continental basins, and lode gold deposits associated with metamorphism. Nevertheless, the flux of hydrothermal fluids responsible for these deposits is minor compared to that caused by circulation due to intrusion of magma¹.

There is agreement that many hydrothermal ore systems derive their thermal energy from magmas^{2,3}, with multiple intrusions necessary to maintain activity for hundreds of thousands of years—the life of some hydrothermal systems⁴. But despite the close association of intrusions with many ore deposits, there is still considerable debate about the extent to which these magmas contribute water, metals, ligands (such as sulphur and chloride) and other components to hydrothermal systems. Many early studies of the stable-isotope composition of hydrothermal minerals indicated a dominance of meteoric water in a variety of types of hydrothermal system⁵, leading some workers to assume that the metals were largely scavenged from the rocks through which these fluids moved⁶. More recent studies, coupled with those of active hydrothermal systems, show that magmatic fluids are commonly present, but that their signatures may be masked or erased owing to later overprinting by large volumes of meteoric water, which represents >95% of the water ultimately convected during cooling of an intrusion in the shallow crust³.

For a hydrothermal ore deposit to form, the flow of metal-bearing fluid must be focused and coupled with a precipitation

mechanism operating in a restricted space, details of which may vary from deposit to deposit. Here we take a broader perspective and examine the evidence for the sources of water, metals and other components in magmas, their related hydrothermal systems and the ore deposits that they form. We focus on the magmatic intrusions and associated hydrothermal systems that occur along convergent plate margins, where oceanic crust is subducted under continental lithosphere, leading to the formation of volcanic arcs. This is the tectonic setting where most magma-related hydrothermal ore deposits are concentrated⁷.

We discuss two broad ore types: intrusion-associated porphyry copper deposits, and shallow-formed, epithermal precious and base-metal deposits. We choose these from the range of hydrothermal ore deposit types (Table 1) because they represent the proximal and distal extremes, respectively, relative to magmatic intrusions. Also, some of their active equivalents—such as volcanic systems with high-temperature fumaroles and acidic springs, and geothermal systems with neutral-pH hot springs and geysers—are known, and have thus been studied. This relationship between ore deposits and active hydrothermal systems has long been surmised⁸, though it is barely a decade since there has been general (though still not universal) acceptance that hydrothermal ore deposits form from systems similar to those now active⁹.

Evidence from active hydrothermal systems, coupled with results from a variety of field and experimental studies, clearly indicates that magmas contribute components such as water, metals and ligands to hydrothermal fluids. Not surprisingly, this evidence wanes as the distance from the intrusion increases, meteoric water becomes dominant, and the fluid salinity and acidity decrease. Even in such distal environments, though, there are indications that discrete, episodic addition of magmatic components in the form of high-pressure vapour may be critical to the ore-forming process.

Association between magmas and ore deposits

The genetic association between magmas and hydrothermal ore deposits is well documented in scores of geological studies of mineralization close to magmatic intrusions^{10,11}. Experimental and modelling studies have shown that the metals and ligands in such deposits could have been mobilized entirely from the associated magmas^{12,13}. Furthermore, the flux of metals measured from some erupting volcanoes indicates that, given sufficient time and a concentration mechanism, degassing magmas can exsolve sufficient metals to create an ore deposit (Table 2). For example, White Island, New Zealand (see cover) is an arc-related volcano with an associated hydrothermal system that has been active for at least the past 10,000 yr. The long-term

TABLE 1 Representative hydrothermal ore deposits associated with subduction-related magmatism

Ore deposit type	Relation to magma	Temperature Depth	Fluid	Associated metals	Example of active analogue
Porphyry	Adjacent to or hosted by intrusion	>600 to 300 °C 2–5 km	Hypersaline and immiscible vapour	Cu ± Mo ± Au, Mo, W or Sn	Shallow magma bodies beneath stratovolcano
Skarn	Adjacent to intrusion in carbonate rock	400–600 °C 1–5 km	Saline to moderately saline	Fe, Cu, Sn, W, Mo, Au, Ag, Pb–Zn	Shallow magma bodies beneath stratovolcano
Pluton-related veins	Fractures in and near intrusion	300–450 °C Variable	Moderate to low salinity	Sn, W, Mo ± Pb–Zn, Cu, Au	Shallow magma bodies beneath stratovolcano
Epithermal (high sulphidation)	Above parent intrusion	<300 °C Near-surface to >1.5 km	Moderate to low salinity, early acidic condensate	Au–Cu Ag–Pb	High-temp. fumaroles and acidic springs near volcanic vent
Epithermal (low-sulphidation)	Distant (?) from magmatic heat source	150–300 °C Near-surface to 1–2 km	Very low salinity, gas-rich, neutral pH	Au(Ag, Pb–Zn)	Geothermal systems with neutral-pH hot springs, mud pools
	Distant (?) from magmatic heat source	150–300 °C Near-surface to 1–2 km	Moderate salinity	Ag–Pb–Zn(Au)	Not observed, transient brine?
Massive sulphide	Near extrusive domes	<300 °C on or near sea floor	Near seawater salinity, gas-rich	Zn–Pb–Ag (Cu or Au)	Back-arc seafloor vents, black smokers

The term 'fluid' is used to refer to non-silicate, aqueous liquid and/or vapour. The salinities (Na, K chloride) of fluids in these environments vary from hypersaline (>50 wt%) to moderate (10–20 wt%), low (<5 wt%) and very low (0.2–0.5 wt%) salinity.

SO₂ flux from White Island volcano, measured by correlation spectrometry, is about 0.13×10^6 tons per year ($t \text{ yr}^{-1}$), and is accompanied by $1.9 \times 10^6 t \text{ yr}^{-1} \text{ H}_2\text{O}$ and $0.44 \times 10^6 t \text{ yr}^{-1} \text{ CO}_2$. Because the measured Cu/S ratio in the aerosols of the ~900 °C ash-laden plume is $\sim 1.7 \times 10^{-3}$ (ref. 14), the calculated Cu flux is $110 t \text{ yr}^{-1}$ (with $>350 \text{ kg yr}^{-1} \text{ Au}$). High concentrations of Cu and other metals such as Au are also observed in the aerosols that accompany other erupting volcanoes, such as Augustine in Alaska, Mt Etna in Italy (Table 2) and elsewhere.

Other evidence bearing on the relationship between magmas and ore deposits comes from some correlations between the type of metal deposit and the composition of the associated magma. For example, peralkaline intrusions with high K₂O or Na₂O can host ores of lithophile elements such as Zr, Nb and lanthanides, whereas more aluminous F-rich systems are typically associated

with Sn, Mo and B ore bodies. Tin, and commonly W, tends to be found associated with reduced magmas (as indicated by the absence of magnetite), whereas Cu and Mo ore deposits tend to occur with more oxidized magmas (containing magnetite)^{7,15,16}. Even within volcanic arcs, correlations exist between the tectonic environment and the metallogeny of an ore district¹⁷. If ore formation were controlled solely by circulating meteoric water, these various correlations would be unexpected.

Sources of metals in magmas

Notwithstanding the evidence cited above for an association between magmas and ore deposits, we still want to know where magmas acquire their metals and how these metals are concentrated to form ore. Metals can enter magmas through a variety of pathways, including mantle melting, mass transfer from the subducting slab and melting of the crust. Most of Earth's siderophile (Fe-loving) elements such as Sn, Mo, Au and the Pt-group elements may be locked away in the Fe–Ni core, causing their crustal abundances to be much less than those of the global bulk. Nevertheless, many of these elements are also present in Fe–Ni sulphides in the upper mantle¹⁸. During partial melting of the mantle, these sulphides are partially consumed, and contribute metals to basaltic liquids that then ascend into the crust, both along mid-ocean ridges¹⁹ and at subduction zones (Fig. 1). In the former setting, these erupted basalts are commonly altered, and together with their associated oceanic sediments, are eventually subducted below the continental crust, inducing melting in the mantle wedge beneath the volcanic arc. Many arc magmas show clear evidence for the addition of components such as oxidized sulphur, alkali-group elements, water and Cl from the subducted slab^{20–22}, transferred as a fluid or a silicate melt. The lack of abundant Cl-bearing phases in the mantle, coupled with the high Cl contents of arc magmas (700–2,000 p.p.m.), further supports the idea that Cl is recycled during subduction of seawater-altered oceanic crust^{20,23}. Subduction may also contribute additional metals, such as Cu and Zn, to the resulting arc magmas, derived from hydrothermal mineralization of the subducted sea floor^{20,24,25}.

Ore components may also be acquired during magma transit through the crust. For example, the isotopic signature of porphyry Cu deposits in Arizona is strongly correlated with that

TABLE 2 Magmatic fluids discharged from volcanoes

	White Island, NZ ¹⁴	Satsuma Iwojima, Japan ⁶⁸	Augustine, Alaska ^{69,77}	Mt Etna, Italy ^{78,79}
Year	1988	1990	1976	1987
Style of discharge	Eruption	Fumarole	Eruption	Fumarole
Temp. (°C)	>850	877	>900	870
Flux ($10^6 t \text{ yr}^{-1}$)				
H ₂ O	1.9	5.2	ND	0.03
CO ₂	0.5	0.04	ND	0.003
Cl	0.04	0.06	>0.6	0.005
S	0.06	0.09	>0.2	0.005
Cu ($t \text{ yr}^{-1}$)	110	0.16	>1200	0.011
Au ($kg \text{ yr}^{-1}$)	>36	0.02	ND	ND
Abundance ratios*				
Na/S ($\times 10^3$)	55	0.08	50–800	0.4
Cu/S ($\times 10^6$)	2,400	1.9	2,300–4,700	2.3

ND, not determined.

* For comparison with the erupted aerosols, the Na/S and Cu/S of fumaroles at White Island are 0.7×10^{-3} and 2.8×10^{-6} , respectively.

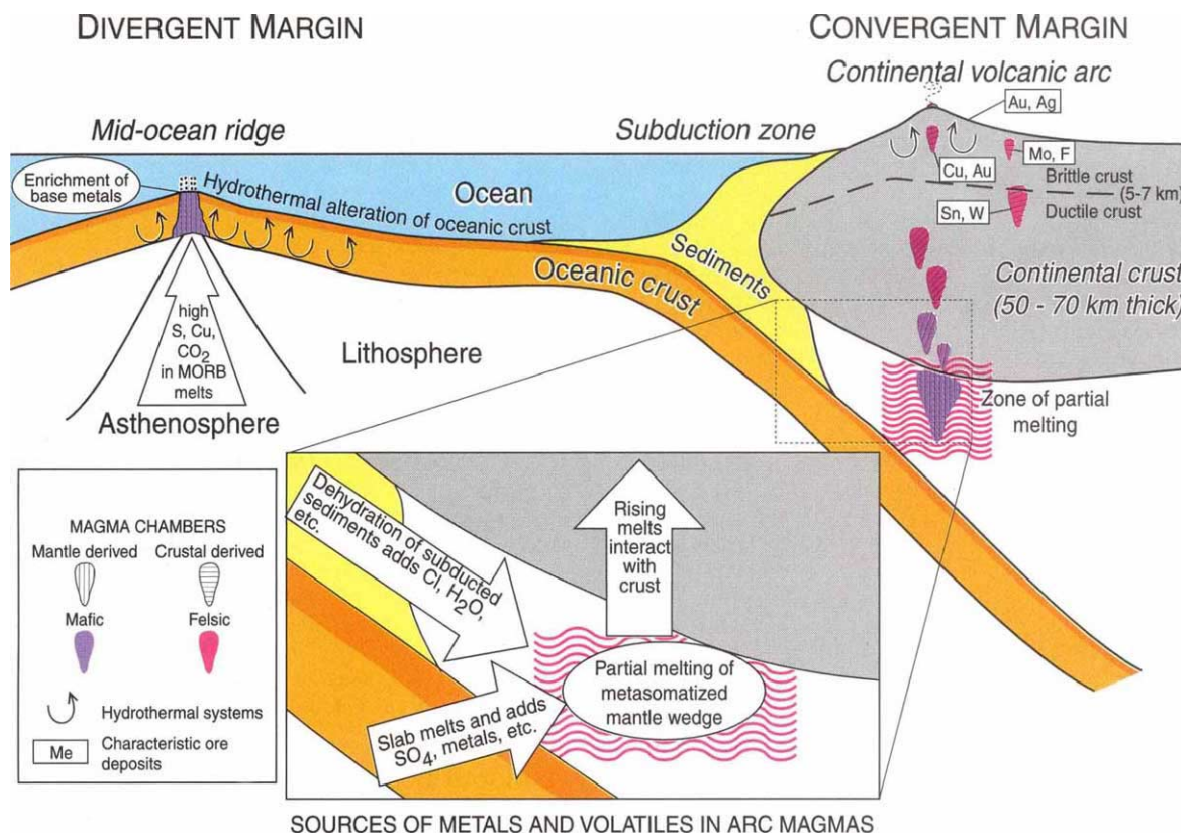


FIG. 1. Schematic section showing the principal components of magma genesis, fluid flow and metallogensis in divergent- and convergent-margin settings.

of the local basement, implying that the intrusions contain a significant amount of assimilated crust²⁶, whereas in the case of Sn and W deposits such as those of southwest England and southwest Japan, the ore-bearing magmas may be produced entirely by melting of the crust¹⁵. Crustal melts produced in the presence of either elemental C or Fe sulphide may be reduced in nature, and are thus ideal for accumulation of Sn¹⁶. Fusion of previously melted crustal rocks can result in high magmatic concentrations of F, due to the breakdown of refractory, F-rich hydrous silicates. These minerals may also contain significant amounts of Sn and Mo, which are concentrated with F in some ore deposits. But exceptions to these relationships are common, and the source of the metals may be less important, in terms of metal concentration, than the magmatic processes that operate during final ascent and differentiation of the magma.

The effects of crystal fractionation. Crystallization is one of the main controls on the concentration of ore components in magmas. The dominant minerals that crystallize in most igneous systems are silicates and oxides; elements that partition into these minerals can have their concentrations lowered, whereas incompatible elements (those that remain in the melt) may become more concentrated. The degree of compatibility is a function of crystal structure, melt composition, magma oxidation state, and the temperature and pressure of crystallization. For example, Mo and Zn concentrations are commonly high in magmas that contain low abundances of Fe-Ti oxides and titanite, minerals that tend to be rich in these elements²⁷. The enrichment of Sn and W in granites formed from melting of crustal rocks, noted above, may also be due to the composition and low abundance of oxide minerals^{28,29} in these reduced rocks. The lack of Cu mineralization associated with high-silica granites may be attributed to the early crystallization of pyrrhotite (Fe_{1-x}S), which acts as a sink for Cu (and Au?) and lowers its concentration in the melt²⁸. Conversely, because Mo starts at a lower concentration and is more incompatible than Cu, crystal fractionation is necessary to produce the higher Mo/Cu of high-

silica granites²⁸. But even after extensive crystallization, metal concentrations in the magma are still at least two or three orders of magnitude less than grades of economic ore in associated porphyry-style mineralization (Table 3). Other processes are therefore necessary to concentrate metals from the magma into an ore deposit.

Fluids in magma. Given the appropriate fluid composition, most ore metals can be partitioned strongly into a magmatic-hydrothermal fluid as chloride, bisulphide and hydroxyacid complexes³⁰⁻³³ (Table 3). At magmatic temperatures and

TABLE 3 Typical metal concentrations in granites and porphyry ores

Metal	Average concentration in granite (p.p.m.) ⁸⁰	Common phenocryst host	Typical concentration (p.p.m.) in porphyry ore ⁸¹	Likely ligand component in magmatic fluid	Maximum vapour/melt partition coefficient ^{28,31,35}
Cu	12	po	>5,000	Cl	100-200
Pb	20	Kf, bi	40-100+	Cl	10-20
Zn	50	FeOx and FeMg	60-100	Cl	15-25
Mo	1.5	FeOx and FeMg	>1,000*	OH	2-10
Sn	3	FeOx and FeMg	3,000*	Cl	>1
W	1.5	FeOx and FeMg	5,000*	OH	4
Ag	0.04	?	1-3 [†]	Cl?	?
Au	0.002	po	<0.1- >0.5 [†]	Cl?(HS)	?

Phenocrysts: the common phenocryst host for the metal in crystallizing granitic plutons. Abbreviations: po, pyrrhotite; FeOx, Fe-Ti oxides; FeMg, ferromagnesian minerals such as fayalite, pyroxene, amphibole and biotite; Kf, K-feldspar; bi, biotite. Likely ligand and common phenocryst host are from various sources, though little is known about either; OH refers to hydroxyacid complexes.

*For Mo-, Sn- or W-rich porphyry, not typical of Cu porphyry ore.

[†]Concentrations listed are for disseminations in porphyry systems. In epithermal vein deposits, Au commonly ranges from 3 to 10 p.p.m., and can exceed 100 p.p.m. as an average in a few deposits.

pressures, insufficient data exist to define the ideal conditions for partitioning of any given metal from a magma into an exsolved hydrothermal fluid. Even so, the limited data available indicate that metals such as Cu^{28,34} can be scavenged by exsolving magmatic fluids, especially at low magmatic pressures³⁵, raising the question: when do these fluids separate from magmas and what controls the fluid composition?

Many magmas contain exsolved fluids during their ascent through the middle to upper crust. There is evidence that basalts from a variety of tectonic settings are saturated with a CO₂-rich fluid phase at depths >30 km (refs 36–38). Because of its high solubility in silicate melt, H₂O will represent only a small fraction of this fluid in the deep crust, although H₂O increasingly exsolves at low pressures until it dominates the fluid, as observed in volcanic emanations (Table 2). Unless felsic magmas originate as isolated systems, and receive no input of volatiles from mafic forerunners and subjacent heat sources, they too are likely to be saturated with a fluid phase throughout their history³⁹. The addition of SO₂ and other gases to magmatic systems decreases the solubility of H₂O and CO₂ in the melt by decreasing their activity in the fluid phase(s), resulting in fluid saturation with even lower concentrations of gases dissolved in the melt. Thus, if magmas in the upper crust are normally fluid-saturated, the key questions related to metal transport are not if or when fluid saturation occurs in the associated magma, but what controls fluid composition, which fluids are most likely to scavenge metals

from the silicate melt, and how much fluid can be channelled to a site of metal deposition.

At shallow crustal depths, metal partitioning becomes more complicated because magma may contain more than one non-silicate fluid phase. For example, at 1 kbar and 800 °C, the NaCl–H₂O system will consist of two phases (a low-density vapour and a hypersaline liquid; Fig. 2) if the bulk composition of the fluid(s) exsolved from the melt is between ~2 and ~70 wt% NaCl. Indeed, the Cl contents of many silicic magmas are consistent with the concentrations necessary to result in saturation with both vapour and a hypersaline liquid^{4,40}. Addition of CO₂ (and most probably SO₂) will increase the field of immiscibility so that the two aqueous phases may coexist over a wide range of crustal conditions (up to pressures ≥5 kbar, ref. 41).

The large density contrast between the vapour and hypersaline liquid commonly results in these two phases separating within the magma chamber or its adjacent hydrothermal system^{4,42}. The aqueous vapour contains CO₂, SO₂, H₂S, HCl and so on, and tends to ascend on formation, either discharging to the surface as volcanic fumaroles (Table 2), or becoming absorbed at depth, forming an acidic water that can leach the host rock⁴². The dense hypersaline liquid remaining at depth will be richer in Cu and other chlorophile elements than the low-salinity vapour³³. Evidence for these immiscible fluids is abundant in intrusive rocks and some volcanic products⁴³. However, it is in porphyry ore deposits that the remnants of these fluids, trapped as inclusions in minerals, have been studied most intensively.

Mineralization associated with intrusions

Porphyry Cu mineralization is dominated by magmatic fluid in the early stage, although late meteoric water is not only common but is perhaps critical in enhancing porphyry metal concentrations to ore grade^{10,33}. These deposits typically consist of sulphides that are both disseminated and located in veinlets; these occurrences are found within and adjacent to porphyritic, silicic to intermediate-composition intrusions, and commonly form at depths of 2–5 km under lithostatic pressures. Detailed studies of hydrothermal alteration, mineralization and fluid inclusions, integrated with stable-isotope data, indicate that the hydrothermal systems that form these deposits are initially hot (>500–600 °C) and dominated by a magmatic, hypersaline liquid coexisting with a low-density vapour. Because of the depth and high temperature, active examples have yet to be studied. Zoned alteration assemblages form as the reactive, metal-bearing magmatic fluid moves away from the intrusion, cools and reacts with the country rock^{10,33}. The intrusions themselves are altered, with a central zone of biotite ± K-feldspar (K-silicate zone), that typically gives way outward to an intermediate zone of quartz plus chlorite, commonly overprinted by sericite and pyrite (sericite zone). The wall rocks farther from the intrusion may be altered to epidote, chlorite and albite (propylitic zone). High-temperature Cu-sulphide mineralization, accompanied by magnetite, may be present as a shell at the boundary of the K-silicate and sericite zones. Where preserved, these alteration zones are always capped by rock that has been leached by an acidic water formed by the absorption of magmatic vapour into meteoric water. The resultant advanced argillic alteration usually extends to the palaeosurface¹⁰.

The strong spatial association between plutons and porphyry-related ore deposits emphasizes the genetic role of magmas in the mineralization process. Another line of evidence comes from the fluids trapped as inclusions in minerals⁴; many of these inclusions formed at near-magmatic temperatures of 700–800 °C. At room temperature, the hypersaline fluid inclusions contain a brine, a vapour bubble and daughter crystals of halite, sylvite and minor ore minerals such as chalcocite (Fig. 2a). The hypersaline liquid at its trapping temperature contains 40 to >60 wt% NaCl equivalent (eq.) and 250–5,000 p.p.m. Cu³⁷, the latter based on the amount of Cu-bearing daughter minerals and recently confirmed by direct analysis⁴⁴. Inclusions that trapped

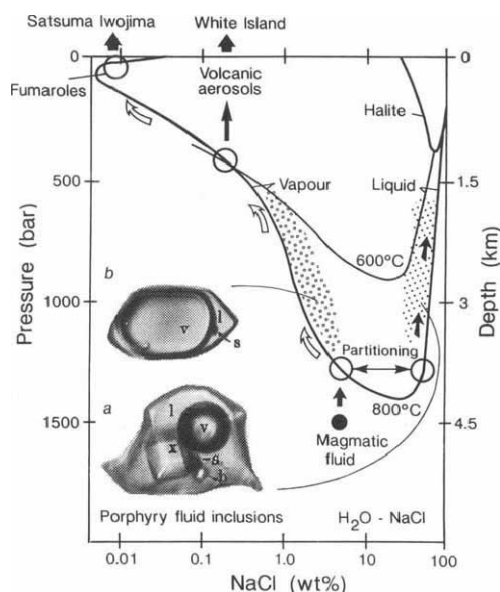


FIG. 2 The composition of immiscible fluids in the system H₂O–NaCl⁸², as a function of pressure at 800 °C (and 600 °C, light line); the approximate depth is for a lithostatic pressure gradient. If a 5 wt% NaCl solution (filled circle) exsolves from a magma at 800 °C and 1,500 bar and then isothermally decompresses, it will unmix to give a low-density vapour and a high-density, hypersaline liquid at ~1,300 bar; chloride-complexed metals will partition to the liquid, whereas gaseous components will be largely contained by the vapour. Upon further isothermal decompression (curved arrows), the liquid and vapour will progressively become more and less saline, respectively, by condensation of NaCl-rich liquid from the vapour. The vapour of high-temperature volcanic fumaroles contains <100 p.p.m. NaCl, and is metal poor, whereas high-pressure vapour accompanying volcanic eruption is relatively NaCl- and metal-rich (for example, White Island⁸³, Table 2). The stippled fields indicate the compositions of fluid inclusions at temperatures and pressures typical of porphyry Cu formation. The photographs show representative quartz-hosted fluid inclusions (at room temperature) from a porphyry Cu deposit (a, hypersaline inclusion containing NaCl-saturated liquid, l; vapour bubble, v; daughter minerals of halite, x; possibly anhydrite, b, and hematite, s; b, vapour-rich inclusion containing low-salinity condensate, l; bubble of compressed gas, v; and small opaque mineral, s). Inclusions are each ~50 µm long.

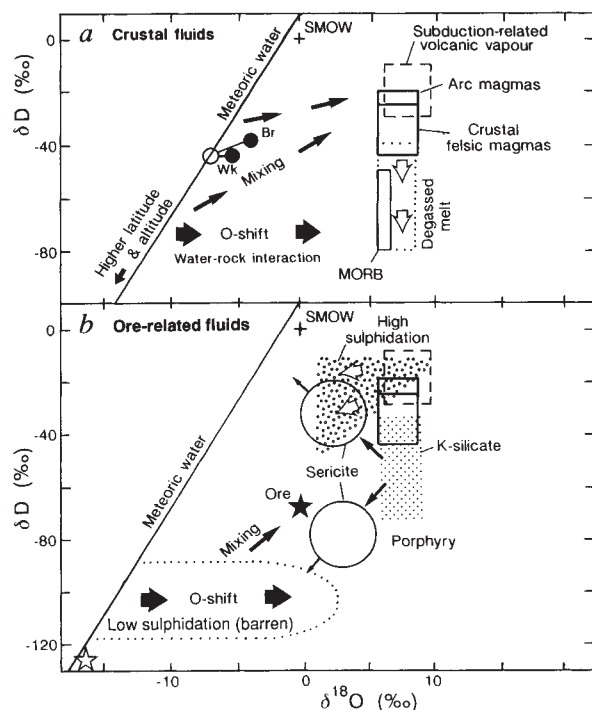


FIG. 3 a, Variation in the O- and H-isotope composition of crustal magmas and waters ($\delta^{18}\text{O}$ and δD , respectively, relative to Standard Mean Ocean Water—SMOW⁵). Felsic magmas underlain by continental crust have slightly heavier H-isotope ratios than island arc magmas, presumably due to incorporation of crustal material⁴⁷. These magmatic waters are both isotopically distinct from that associated with mantle-derived mid-ocean-ridge basalts (MORB)⁴⁶. The isotopic composition of 'subduction-related magmatic vapours' outgassing from arc volcanoes has a narrow range²¹, enriched from their parent arc magmas due to fractionation during degassing⁴⁶, with the degassed melt (open arrows pointing down) becoming lighter in H-isotope composition. There are two trends for neutral-pH geothermal waters, one caused by meteoric water-rock interaction ('O-shift' from the meteoric water line⁶⁴; for example, that of Wairakei, Wk), and the other from mixing between meteoric water of variable (local) isotopic composition and low-salinity magmatic vapour²¹ (small solid arrows with trends pointing towards volcanic vapour; for example, that of Broadlands, Br). Connate waters trapped in sediments and waters related to metamorphism have wide ranges of isotopic composition⁸⁵ (not shown here). b, The isotopic composition of hydrothermal fluids associated with porphyry ore deposits is deduced from the composition of high-temperature K-silicate alteration minerals (stippled field)^{5,45}. Lower-temperature sericitic alteration indicates various mixtures (typically plotting in the two large circles) between late-stage magmatic water (derived from a degassed melt⁴⁸) and meteoric water with a range of isotopic composition depending on the palaeolatitude and altitude; this meteoric water is sometimes O-shifted⁴⁹. 'High-sulphidation' waters⁵⁴ (see text) are similar in isotopic composition to volcanic vapour, with trends projecting towards meteoric water of variable isotopic composition (open arrows) depending on location of the deposit. The waters that precipitate barren quartz in 'low-sulphidation' deposits commonly show an O-shift from local meteoric water values^{5,61}, whereas high-grade ore samples (Comstock lode; solid star) have both an O- and H-isotopic shift from local meteoric water (open star) caused by a component of magmatic water⁶¹; such a huge shift cannot be caused by boiling of meteoric water²¹.

the immiscible vapour (Fig. 2b) are low in salinity (<1–2 wt% NaCl eq.) and relatively high in CO_2 (≤ 10 wt%; ref. 37).

Constraints from O and H isotopes. The isotopic composition of water in a magma was initially approximated by analysing minerals in eroded intrusions and then using mineral–water fractionation factors³ to calculate the values of the coexisting water. These data were used to define the isotopic composition of 'primary magmatic water'⁴⁵ (Fig. 3a; dotted-line box). But because there is a strong fractionation of H isotopes during early

degassing⁴⁶, the late-formed hydroxyl-bearing igneous minerals represent the isotopic composition of a degassed melt rather than that of the initial magmatic water (Fig. 3a; open arrows). The best current estimate of the initial isotopic composition of H_2O in arc and crustal felsic magmas⁴⁷ (Fig. 3a; solid-line boxes) is intermediate between that of the degassed magma (dotted-line box) and the isotopically heavy outgassed H_2O , the latter sampled as volcanic vapour²¹ (dashed-line box).

The O- and H-isotope compositions of hydrothermal minerals in ore deposits also help to trace the origin(s) of the hydrothermal water. In porphyry deposits, hydroxyl-bearing minerals include biotite, which forms close to the intrusion at near-magmatic temperatures (>500–600 °C), chlorite and sericite (which form at ~300 °C) and kaolinite (~200 °C). Early-formed, high-temperature biotite and K-feldspar^{5,45} provide evidence that the $\delta^{18}\text{O}$ of the fluids altering the intrusions have a narrow range of +6 to +9‰ (Fig. 3b; stippled box), consistent with the water being in equilibrium with the large O reservoir in the associated magmas. But the δD value of this early magmatic fluid (based on biotite compositions) has a wide range, from –35 to –75‰ (Fig. 3b), and is shifted to lower values than the primary water of most felsic and arc magmas (–20 to –45‰). This may be explained by open-system degassing of the magma⁴⁶, as noted above, which can thus account for the large δD variations (>40‰) at constant O-isotope composition observed for biotite from a single deposit⁴⁸.

Lower-temperature alteration commonly has an isotopic signature indicating a large component of meteoric water^{45,49}, although in some deposits there is also evidence that sea water or an evolved formation water was present, and this saline liquid may have had a role in mineralization⁴⁸. The isotopic variations of lower-temperature, sericitic alteration among several North American deposits are systematic with latitude and hence palaeo-meteoric water composition⁴⁵ (Fig. 3b; hydrothermal waters typically plotting in the circled areas). This meteoric water has commonly been shifted in O-isotope composition through water-rock interaction before its flow past the intrusion⁴⁹ (for example, lower circled area). But in some South American deposits⁵⁰, there is clear evidence that the fluid related to late-stage sericite samples had a magmatic vapour component, as their compositions indicate a fluid with +5 to +8‰ $\delta^{18}\text{O}$ and –20 to –40‰ δD (Fig. 3b; dashed-line box).

Other isotopes. The Pb- and S-isotope compositions of porphyry Cu ores from South and North American porphyry deposits^{51,52} are equivocal regarding their origin. In some cases there are Pb-isotope trends consistent with subducted pelagic sediments as the major source of Pb, indicating that Pb was recycled during subduction and then contributed to the porphyry hydrothermal system. Crustal sources appear only locally important to the Pb budget of most of these systems; in contrast, Sr sometimes shows high isotopic ratios, consistent with a large crustal component⁵¹. In Arizona, however, there is a strong correlation between mineralizing intrusions and the Pb-isotope composition of the lower-crustal basement²⁶. The range of $\delta^{34}\text{S}$ of most porphyry Cu deposits lies between –3 and +9‰ (ref. 53), suggesting the bulk of the S was derived from local intrusions, though the wide range of S-isotope composition in some cases may be caused by contribution from the wall rock^{50,53}. The fresh-rock $\delta^{34}\text{S}$ value has a geographical variation, which in island arcs probably reflects a contribution from subducted marine sediments⁵³. Thus isotopic evidence from porphyry deposits indicates a variety of sources for components such as H_2O , Pb and S, although it seems that a magmatic source dominates the early history of the systems. As the distance from the magma to ore increases, such evidence weakens.

Mineralization distant from magmatic intrusions

The term 'epithermal' refers to a class of hydrothermal ore deposits formed at relatively low temperatures (<300 °C) and shallow depths (surface to 1–2 km; Table 1). Despite the shallow

depth, there is still evidence for a magmatic component to the fluids.

Epithermal ore deposits can be divided into two main types (Table 1), according to their alteration mineralogy. One type is formed by acidic and oxidized fluid (high sulphidation), typical of acidic springs near volcanoes; the other type is formed by a near-neutral pH, reduced fluid (low sulphidation), as found in geothermal systems⁹.

High-sulphidation epithermal ore deposits. Although formed at shallow depths, high-sulphidation ore deposits have characteristics that indicate involvement of magmatic components in their formation⁵⁴. The early stage of these deposits is characterized by extensive leaching of the host rocks by fluids with a pH <2 (ref. 55) and an O- and H-isotope composition (Fig. 3b; triangular area)⁵⁴ similar to that of magmatic vapour²¹ mixed with meteoric water (Fig. 3b; open arrows). The leaching creates a porous silica residue (>95 wt% SiO₂) that may subsequently host ore minerals such as Au-bearing Cu- and Fe-sulphides. The high-sulphidation style of mineralization shares many mineralogical and stable-isotope characteristics with the advanced argillic zone of alteration that caps porphyry Cu deposits^{33,54}, and indeed there is commonly a close spatial relationship between these deposits¹¹. The hydrothermal fluid is oxidized, based on S-isotope compositions of the sulphide and sulphate minerals⁵⁴; a majority of the sulphur species originate from the disproportionation of magmatic SO₂ to form H₂SO₄.

Fluid inclusions in ore-related minerals, including Cu-sulphosalts (transparent only to infrared wavelengths) and quartz, indicate that the mineralizing fluid may have a salinity of 2–5 wt% NaCl eq., though there can be a marked increase in salinity with depth, to 20–30 wt% NaCl eq. below the ore zone⁵⁶. This spatial separation of low-salinity liquid associated with high-sulphidation ore above higher-salinity liquid at depth is distinct from that noted for deeper porphyry Cu deposits, where hypersaline brine coexists with low-salinity vapour; this separation may be caused by the high density of the brine, which cannot ascend easily to shallow depths. The close association of Cu and Au in the relatively oxidized high-sulphidation and porphyry environments suggests that both metals may be transported under similar conditions, though in the high-sulphidation epithermal environment mineralization is related to a lower temperature and less saline liquid.

Low-sulphidation epithermal ore deposits. In contrast to porphyry and high-sulphidation deposits, magmatic signatures in low-sulphidation epithermal Au deposits are more elusive. Ore concentrations of Cu do not occur, presumably because the low-salinity, near-neutral pH and reducing (sulphide-dominant) fluids preclude efficient transport of Cu. In general, low-sulphidation deposits form distant from the inferred magmatic heat source¹¹ at temperatures of 200–300 °C (ref. 57). Pressures are controlled by hydrostatic conditions, meaning that the maximum temperature at a given depth is constrained by boiling, a common process in these systems⁵⁸. Although a variety of classifications of this deposit type exist⁵⁹, a division based on metal assemblage appears to reflect intrinsic geochemical differences. One style is Au-rich (with Ag/Au of 1/10 to 10/1 and only trace base metals), whereas the other is Ag-rich (with Ag/Au of >100/1) with economic quantities of Zn and Pb present. The Au-rich deposits are associated with low-salinity but gas-rich fluids (<1–2 wt% NaCl but up to 4 wt% gases, mainly CO₂ with H₂S)⁵⁸. By contrast, deposits rich in Ag and base metals are associated with more saline fluids (10–15 wt% NaCl)⁶⁰. This difference in salinity is important for the metal-transporting capacity of the fluids, as Au is transported as a bisulphide complex whereas Ag, Zn and perhaps Pb are dependent on transport by Cl complexes in this reduced environment³². The low-salinity fluids of Au-rich epithermal deposits are largely similar to the fluids of most active geothermal systems^{9,58}, whereas there are no saline geothermal systems in a tectonic environment similar to that hosting Ag- and base metal-rich epithermal deposits.

This lack may be due to transient, tectonically-induced pulses of saline, mineralizing fluid into an otherwise low-salinity, metal-barren hydrothermal system⁶⁰, a process we have not yet seen.

Stable-isotope studies of low-sulphidation deposits commonly demonstrate the predominance of meteoric water⁶¹, with O-isotope compositions of both water (Fig. 3b; 'O-shift') and rock modified to varied degrees by their mutual interaction⁵. However, most samples studied to date have been late-stage, barren quartz veins, formed after the period of ore deposition⁶¹. Where samples are associated directly with bonanza (high-grade) ore, such as from the famous Comstock lode, Nevada (Fig. 3b; solid star), a large shift in both $\delta^{18}\text{O}$ and δD values from the composition of local meteoric water (open star) is taken as evidence for a significant magmatic-water component⁶¹. Paradoxically, the evidence for meteoric-water dominance in epithermal systems is most obvious where magmatism was most voluminous, because the large magmatic heat source creates large and long-lived cells of convecting meteoric water, convection that will erase most of the evidence for any early component of magmatic water. For example, in the Pacific Northwest of North America, a region of many eroded plutons, the rocks exposed over >5% of the region bear an O-isotope signature indicating alteration by meteoric water⁶². Furthermore, the S- and C-isotope evidence indicates that many epithermal deposits can have a magmatic source for both elements⁵³, although there is commonly also a sedimentary component to the S, as well as occasionally an influx of organic C. The results of Pb-isotope studies on a variety of epithermal deposits indicate several sources for the Pb, including leaching of basement rocks and the overlying volcanic rocks or their intrusive equivalents⁶³, not surprising given that meteoric-water convection can extend to depths of 5–7 km (ref. 62).

The multiple sources of components in low-sulphidation epithermal fluids is caused largely by the extensive amount of interaction between thermal waters and crustal rocks that occurs in this environment. Any reactive magmatic components are neutralized and reduced, in turn deriving components from the host rocks. This process leads to an integrated isotopic signature that may mask a variety of discrete fluids present in the early stages of activity; that is, pulses of fluid recorded by cycles of mineral deposition. Despite the overprinting that may occur in epithermal deposits, advances in microanalytical techniques now allow the isotopic composition of preserved domains in minerals to be analysed, domains that show evidence for a variety of discrete water and gas sources⁶⁴, including magmatic and crustal components⁶⁵.

Geothermal systems. In some volcanic arc-hosted geothermal systems, the evidence for a magmatic component is common and unequivocal. These systems are characterized at depth by neutral-pH fluids of low salinity (<1 wt% NaCl) but variable gas content (up to 4 wt%, including CO₂ and H₂S)⁵⁸. Coupled with their alteration mineralogy, these characteristics indicate they are analogous to the epithermal systems responsible for forming low-sulphidation ore deposits; indeed, in two of the world's largest and youngest epithermal Au deposits (Hishikari, Japan and Ladolam, Papua New Guinea) late-stage waters (70 to >150 °C) still flow through the veins.

The advantage in studying active systems is that discrete fluids can be sampled, allowing an examination of 'time-slices' through the history of hydrothermal activity. For example, the isotopic compositions of the highest-temperature waters in a variety of these geothermal systems have two different trends (Fig. 3a); one trend suggests a mixture of up to 50% magmatic water with local meteoric water²¹, whereas the other indicates only meteoric water-rock interaction⁵ (horizontal arrows). In the case of the Broadlands geothermal system, New Zealand, the deep thermal water (Fig. 3a; filled circle, Br) has a H-shift that indicates about a 20% component of low-salinity magmatic vapour mixed with local meteoric water (open circle). The deep liquid is also gas-rich (~2 wt%, mainly CO₂), and has a N₂/Ar ratio characteristic of arc magmatism (caused by subduction of N₂-rich

sediments)⁶⁵. By contrast, the deep liquid from the nearby Wairakei geothermal system (filled circle, Wk) has an isotopic composition similar to meteoric water that is slightly O-shifted. This fluid is low in total gas content (<0.1 wt%), and has a N₂/Ar ratio typical of air-saturated ground water, indicating no significant magmatic contribution to the system⁶⁵. This distinction in fluid composition may help explain the observed differences in mineralization between these two systems. The Broadlands system is mineralized locally with Au-rich precipitates in hot springs (to 80 p.p.m.) and wellhead pipes (to 6 wt%, ref. 66), and Pb–Zn sulphides are common at depth, whereas precious metals and Pb–Zn sulphides are absent at Wairakei. The difference in mineralization also correlates with Pb-isotope signatures. Broadlands' sulphide minerals have a homogenous signature, identical to that of local calc-alkaline intrusions, suggesting these are the source rocks, either through contribution from a magmatic fluid or by leaching, whereas the Pb-isotope compositions of altered rocks from Wairakei indicate that most of the Pb was leached from sedimentary rocks.

Similar geothermal systems with variable magmatic inputs (based on chemical and isotopic signatures) are common also in other circum-Pacific volcanic arcs^{65,67}. Drilling in some systems in Japan and the Philippines has encountered deep fractures containing hot, acidic fluids within otherwise neutral-pH waters; the acidity is caused by an absorbed magmatic vapour that is not neutralized by water–rock interaction⁶⁷. Perhaps such a magmatic fluid, although neutralized at depth, directly contributes metals to the Broadlands system. Alternatively, the acidic vapour may contribute only ligands such as Cl[−] and HS[−], causing the resulting fluid to leach metals from the wall rock. This variability in magmatic signatures between geothermal systems may be caused by intermittent contributions (or lack thereof) of magmatic fluid. Thus, gas and isotopic analysis of fluid inclusions in ore-related versus barren minerals will help to determine the role of magmatic contributions to hydrothermal fluid composition and ore formation.

Magmatic fluids and hydrothermal environments

We have mentioned the relationship between magma type and style of intrusion-related mineralization, stressing the need for metal scavenging from the magma by an exsolving aqueous fluid. Here we examine how this magmatic fluid may be involved in mineralization away from the intrusion.

The acidic environment. Studies of high-sulphidation epithermal ore deposits have confirmed the involvement of fluids with isotopic composition similar to high-temperature volcanic vapours^{21,54,68} (Fig. 3b). When such vapours are absorbed by meteoric water, they form acidic springs, which commonly have a pH of 0.5–1.5 (ref. 68). The composition of these acidic springs reflects isothermal dissolution of the rocks through which they flowed, a process that forms leached zones similar to those that host ore in high-sulphidation deposits. However, studies of the metal concentrations of high-temperature fumaroles^{68,69} associated with quiescent degassing of volcanoes indicate that these vapours cannot form ore deposits⁶⁸. For reasonable lifespans of magma degassing and hydrothermal activity related to a single intrusion (~10⁴ yr; ref. 2), the metal fluxes related to fumaroles (Table 2) and their associated acidic springs⁶⁸ are insufficient to account for the amount of Au (10–100 t) and Cu (>10³ t) found in high-sulphidation ore deposits¹¹.

There are several possible explanations for this seeming paradox. First, these deposits may form over much longer time spans, in association with multiple intrusions (as suggested for some porphyry deposits⁴). Second, the acidic water may leach metals as it reacts with the wall rock at depth⁷⁰. Third, the greater depths of these deposits (in contrast to the surficial fumaroles) means that pressures are high, thus increasing the concentration of metal chlorides in the vapour⁷¹. For example, aerosols associated with magmatic eruption are responsible for much higher metal fluxes than quiescently degassing volcanoes, even though

the discharge temperatures and flux of species such as H₂O, S and Cl are similar (compare White Island and Satsuma Iwojima; Table 2). This observation can be explained if a relatively NaCl- and metal-rich vapour, equilibrated with a magma at a pressure of several hundreds of bar (Fig. 2), accompanies eruption, thus accounting for the high Na/S and Cu/S in aerosols, in marked contrast to the low Na/S and Cu/S in the metal-poor fumarolic vapour associated with quiescent degassing. Absorption of a similar high-pressure and metal-rich vapour by a hydrothermal system may therefore form the low-salinity fluid (<5 wt% NaCl) responsible for mineralization in high-sulphidation systems⁶⁸. The deeper, high-salinity (≤30 wt% NaCl) fluid beneath these deposits, referred to earlier, may be related to the immiscible hypersaline counterpart of the vapour.

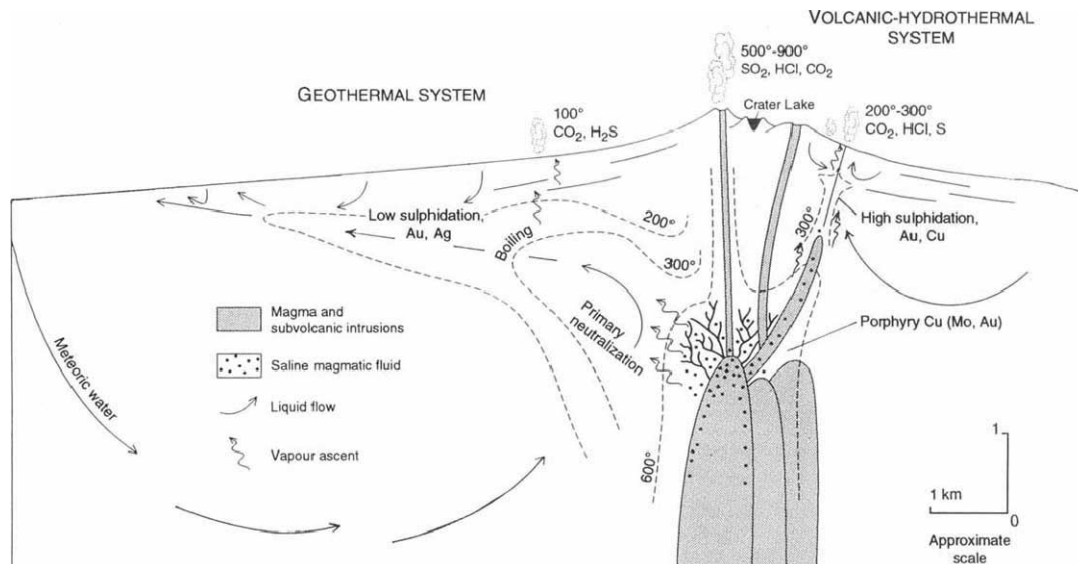
The neutral-pH environment. The source(s) of Au and other metals in low-salinity, neutral-pH fluids of low-sulphidation deposits is open to more speculation compared with porphyry and high-sulphidation systems, where there is evidence for a large magmatic component to the system. If the isotopic signature of water in a low-sulphidation system indicates that the fluid has a component of magmatic water mixed with meteoric water (for example, Comstock, Fig. 3b), then the magmatic component must have a low salinity to account for the dilute waters associated with epithermal Au deposition. Therefore, the magmatic fluid phase must be the low-density vapour noted earlier from volcanic aerosol studies (rather than the hypersaline liquid), which may have an appreciable metal content at the pressures of several kilometres depth⁷¹. Any acidic gases contributed by magmatic vapour to the meteoric-water system will undergo neutralization by reaction with the host rock (Fig. 4), causing an exchange of H⁺ for cations, including metals⁷⁰, thus adding this component to the metals contributed by the vapour. At present, there is no geological evidence that low-sulphidation deposits are underlain by pervasive acidic alteration, suggesting that when magmatic vapours are absorbed at the base of the low-sulphidation epithermal environment, the resulting neutralization may occur over a broad scale⁶⁴ (Fig. 4). This is consistent with the huge expanses (hundreds to thousands of square kilometres) of propylitic alteration (weak hydrogen metasomatism) characteristic of eroded, volcanic-hosted epithermal districts.

We infer the geochemical evolution of a hydrothermal fluid and its resultant alteration and mineralization as the distance from a magma body increases, because of the progressive reaction of fluid with rock. But there is no requirement that a porphyry Cu deposit is overlain by a high-sulphidation deposit, or vice versa, or that a low-sulphidation deposit will be rooted by porphyry mineralization adjacent to the intrusive heat source (Fig. 4). Nevertheless, these are possibilities, and some deposits may reflect this proposed geochemical transition, both in space¹¹ and/or in time. At Porgera, Papua New Guinea, early hypersaline magmatic fluids gave way to later meteoric waters of lower temperature and salinity, with Au probably remobilized from the initial porphyry mineralization⁷². The nearby Ladolam deposit, containing >400 t of Au, had a similar evolution, though the transition from porphyry to epithermal was probably caused by the catastrophic collapse of the host stratovolcano about 300,000 yr ago¹¹. The Kelian low-sulphidation Au deposit in Borneo⁷³ appears to span the epithermal to near-magmatic environment in terms of fluid salinity, temperature and isotopic characteristics. These few examples indicate that it is at present unwise to place sharp boundaries around the types of hydrothermal ore deposits discussed here.

Future directions

The occurrence of ore deposits is uncommon compared to the widespread evidence for extinct hydrothermal activity, indicating that only a small proportion of hydrothermal systems form ore, now or in the past. An outstanding problem in the study of hydrothermal ore deposits is to identify the factors that distinguish systems with ore-forming potential—factors so uncommon

FIG. 4 Schematic cross-section showing shallow, sub-volcanic intrusions, an associated stratovolcano, and the environments deduced for the formation of porphyry Cu, and high- and low-sulphidation epithermal ore deposits^{9,86}. Active volcanic-hydrothermal systems extend from the degassing magma to fumaroles and acidic springs, and incorporate the porphyry and/or high-sulphidation ore environments, whereas low-sulphidation ore deposits form from geothermal systems characterized by neutral-pH waters that may discharge as hot springs and geysers, such as those in Yellowstone National Park, USA.



that perhaps we have yet to observe them in active systems. This perceived lack of active ore-forming systems might also be accounted for by the common evidence indicating the episodic nature of formation of hydrothermal deposits. The fluids responsible for mineralization may be present only for short periods during the lifetime of the hydrothermal system, possibly at times of individual tectonic or hydraulic fracturing events^{4,54,60,74}.

Meteoric waters continue their deep circulation long after an intrusion has solidified. This leads to the masking of signatures of early events in the lives of hydrothermal systems. Nevertheless, advances in stable-isotope microanalysis now allow discrimination among the many events that occur during the histories of hydrothermal systems. Each event may have a different signature for O, H, S and C that can be determined only through study of preserved domains of individual populations of crystals, or even growth zones within crystals, now possible with laser microprobe techniques⁷⁵. Recognition of gas signatures in active systems⁶⁵ can also be applied to the extinct environment because of continuing advances in the analysis of inclusion fluids^{37,64}. Moreover, techniques such as proton-induced X-ray emission (PIXE)⁷⁶ or the X-ray microprobe³⁴ now permit investigations of p.p.m.-level metal concentrations in fluid inclusions as small as several micrometres. By determining the metal contents of a wide range of crustal fluids trapped in minerals, we can better understand why some hydrothermal fluids form ore deposits whereas others do not. As an example, the inclusion fluids associated with granite-hosted Sn mineralization in Australia have been analysed by PIXE⁷⁶. Vapour-rich inclusions appear to have an order of magnitude higher Cu concentrations than coexisting hypersaline inclusions, opposite to the findings in the more oxidized porphyry Cu deposits^{37,44}. Such results provide more questions than answers, and indicate the exciting work ahead on hydrothermal fluids, not only regarding mineralization and the diverse sources of metals, but also the effect these fluids have on crustal processes⁷⁴. □

Jeffrey W. Hedenquist is at the Mineral Resources Department, Geological Survey of Japan, Tsukuba 305, Japan. Jacob B. Lowenstern is at the US Geological Survey, MS 910, 345 Middlefield Road, Menlo Park, California 94025, USA.

1. Cathies, L. M. *Science* **248**, 323-329 (1991).
2. Cathies, L. M. in *Econ. Geol.* 75th Anniv. vol. 424-457 (1981).
3. Norton, D. in *Advances in Geology of the Porphyry Copper Deposits* (ed. Titley, S. R.) 59-72 (Univ. Arizona Press, Tucson, 1982).
4. Fournier, R. O. *Volcanism in Hawaii* (eds Decker, R. W., Wright, T. L. & Stauffer, P. H.) 1487-1506 (Prof. Pap. No. 1350, US Geol. Surv., Washington DC, 1987).

5. Taylor, H. P. in *Geochemistry of Hydrothermal Ore Deposits* 2nd edn (ed. Barnes, H. L.) 236-277 (Wiley, New York, 1979).
6. Ohmoto, H. in *Stable Isotopes in High Temperature Geological Processes* (ed. Valley, J. W., Taylor, H. P. & O'Neill, J. R.) 491-555 (Miner. Soc. Am., Washington DC, 1986).
7. Sawkins, F. J. *Metal Deposits in Relation to Plate Tectonics* 2nd edn (Springer-Verlag, Berlin, 1990).
8. White, D. E. *Econ. Geol.* 50th Anniv. vol. 99-154 (1955).
9. Henley, R. W. & Ellis, A. J. *Earth Sci. Rev.* **19**, 1-50 (1983).
10. Titley, S. R. & Beane, R. E. *Econ. Geol.* **75**, 214-269 (1981).
11. Sillitoe, R. H. *Econ. Geol. Monogr.* **6**, 274-291 (1989).
12. Burnham, C. W. in *Geochemistry of Hydrothermal Ore Deposits* 2nd edn (ed. Barnes, H. L.) 71-136 (Wiley, New York, 1979).
13. Cline, J. S. & Bodnar, R. J. *J. geophys. Res.* **96**, 8113-8126 (1991).
14. Le Cloarec, M. F., Allard, P., Ardouin, B., Giggenbach, W. F. & Sheppard, D. S. *Earth planet. Sci. Lett.* **108**, 19-28 (1992).
15. Ishihara, S. *Econ. Geol.* 75th Anniv. vol. 458-484 (1981).
16. Blevin, D. L. & Chappell, B. W. *Trans. R. Soc. Edinb. Earth Sci.* **83**, 305-316 (1992).
17. Sillitoe, R. H. in *Relations of Tectonics to Ore Deposits in the Southern Cordillera* (eds Dickinson, W. R. & Payne, W. D.) 49-69 (Arizona Geol. Soc., Tucson, 1981).
18. Lorand, J. P. *Geochim. cosmochim. Acta* **54**, 1487-1492 (1990).
19. MacLean, W. H. *Econ. Geol.* **64**, 865-884 (1969).
20. Gill, J. *Orogenic Andesites and Plate Tectonics* (Springer, Heidelberg, 1980).
21. Giggenbach, W. F. *Earth planet. Sci. Lett.* **113**, 495-510 (1992).
22. Albarede, F. & Michard, A. in *Crust/Mantle Recycling at Convergence Zones* (eds Hart, S. R. & Gülen, L.) 29-36 (Kluwer, Dordrecht, 1989).
23. Ito, E., Harris, D. M. & Anderson, A. T. *Jr Geochim. cosmochim. Acta* **47**, 1613-1624 (1983).
24. Sillitoe, R. H. *Econ. Geol.* **67**, 184-197 (1972).
25. Stolper, E. M. & Newman, S. *Earth planet. Sci. Lett.* **121**, 293-325 (1994).
26. Bouse, R. M., Ruiz, J., Titley, S. R., Lang, J. R. & Wooden, J. L. abstr. *Eos* **71**, 1681 (AGU Fall mtg, 1990).
27. Lowenstern, J. B., Mahood, G. A., Hervig, R. L. & Sparks, J. *Contr. Miner. Petrol.* **114**, 119-129 (1993).
28. Candela, P. A. in *Ore Deposits Associated with Magmas* (eds Whitney, J. A. & Naldrett, A. J.) 203-232 (Rev. Econ. Geol. Vol. 4, Soc. Econ. Geol., Littleton, Colorado, 1989).
29. Lehmann, B. *Econ. Geol.* **77**, 50-59 (1982).
30. Holland, H. D. *Econ. Geol.* **67**, 281-301 (1972).
31. Keppeler, H. & Wyllie, P. J. *Contr. Miner. Petrol.* **109**, 139-150 (1991).
32. Seward, T. M. in *Physics and Chemistry of the Earth* (eds Wickman, F. & Rickard, D.) 113-129 (Pergamon, Oxford, 1981).
33. Hemley, J. J. & Hunt, J. P. *Econ. Geol.* **87**, 23-43 (1992).
34. Lowenstern, J. B., Mahood, G. A., Rivers, M. L. & Sutton, S. R. *Science* **252**, 1405-1409 (1991).
35. Urabe, T. *Econ. Geol.* **82**, 1049-1052 (1987).
36. Gerlach, T. M. & Taylor, B. E. *Geochim. cosmochim. Acta* **54**, 2051-2058 (1990).
37. Roedder, E. *Fluid Inclusions* (Miner. Soc. Am., Washington DC, 1984).
38. Javoy, M. & Pineau, F. *Earth. planet. Sci. Lett.* **107**, 598-611 (1991).
39. Holloway, J. *Geol. Soc. Am. Bull.* **87**, 1513-1518 (1976).
40. Lowenstern, J. B. *Am. Miner.* **79**, 353-369 (1994).
41. Bodnar, R. J. & Sterner, S. M. in *Hydrothermal Experimental Techniques* (eds Ulmer, G. C. & Barnes, H. L.) 423-457 (Wiley, New York, 1987).
42. Henley, R. W. & McNabb, A. *Econ. Geol.* **73**, 1-20 (1978).
43. Roedder, E. *Geochim. cosmochim. Acta* **56**, 5-20 (1992).
44. Bodnar, R. J. et al. abstr. *Eos suppl.* **74**, 669 (AGU Fall mtg, 1993).
45. Sheppard, S. M. F., Nielsen, R. L. & Taylor, H. P. *Jr Econ. Geol.* **66**, 515-542 (1971).
46. Taylor, B. E. in *Stable Isotopes in High Temperature Geological Processes* (eds Valley, J. W., Taylor, H. P. & O'Neill, J. R.) 185-225 (Miner. Soc. Am., Washington DC, 1986).
47. Taylor, B. E. in *Magmatic Contributions to Hydrothermal Systems* (ed. Hedenquist, J. W.) 190-194 (Rep. No. 279, Geol. Surv. Japan, Tsukuba, 1992).
48. Dilles, J. H., Solomon, G. C., Taylor, H. P. & Einaudi, M. T. *Econ. Geol.* **87**, 44-63 (1992).
49. Zaluski, G., Nesbitt, B. E. & Muehlenbachs, K. *Econ. Geol.* (in the press).
50. Kusakabe, M., Hori, M. & Matsuhisa, Y. in *Stable Isotopes and Fluid Processes in Mineralization* (eds Herbert, H. K. & Ho, S. E.) 244-259 (Univ. W. Australia, Perth, 1990).
51. Sillitoe, R. H. & Hart, S. R. *Geochim. cosmochim. Acta* **48**, 2135-2142 (1984).

52. Church, S. E., LeHuray, A. P., Grant, A. R., Delevaux, M. H. & Gray, J. E. *Geochim. cosmochim. Acta* **50**, 317–328 (1986).
53. Ohmoto, H. & Rye, R. O. in *Geochemistry of Hydrothermal Ore Deposits* 2nd edn (ed. Barnes, H. L.) 509–567 (Wiley, New York, 1979).
54. Rye, R. O. *Econ. Geol.* **88**, 733–753 (1993).
55. Stoffregen, R. E. *Econ. Geol.* **82**, 1575–1591 (1987).
56. Sanger-von Oepen, P., Friedrich, G. & Kisters, A. *Miner. Deposita* **25**, S36–S41 (1990).
57. Heald, P., Foley, N. K. & Hayba, D. O. *Econ. Geol.* **82**, 1–26 (1987).
58. Hedenquist, J. W. & Henley, R. W. *Econ. Geol.* **80**, 1379–1406 (1985).
59. White, N. C. & Hedenquist, J. W. *J. Geochim. Explor.* **36**, 445–474 (1990).
60. Simmons, S. F., Gemmell, J. B. & Sawkins, F. J. *Econ. Geol.* **83**, 1619–1641 (1988).
61. O'Neil, J. R. & Silberman, M. L. *Econ. Geol.* **69**, 902–909 (1974).
62. Criss, R. E., Fleck, R. J. & Taylor, H. P. Jr *J. geophys. Res.* **96**, 13335–13356 (1991).
63. Doe, B. R. & Zartman, R. E. in *Geochemistry of Hydrothermal Ore Deposits* 2nd edn (ed. Barnes, H. L.) 22–70 (Wiley, New York, 1979).
64. Norman, D. I. & Musgrave, J. A. *Geochim. cosmochim. Acta* **58**, 1119–1131 (1994).
65. Giggenbach, W. F. *Econ. Geol.* **87**, 1927–1944 (1992).
66. Brown, K. L. *Econ. Geol.* **81**, 979–1983 (1986).
67. Reyes, A. G., Giggenbach, W. F., Saleras, J. R. M., Salonga, N. D. & Vergara, M. C. *Geothermics* **22**, 479–519 (1993).
68. Hedenquist, J. W., Aoki, M. & Shinohara, H. *Geology* **22**, 585–588 (1994).
69. Symonds, S. B., Reed, M. H. & Rose, W. I. *Geochim. cosmochim. Acta* **56**, 633–657 (1992).
70. Reed, M. H. in *Magmatic Contributions to Hydrothermal Systems* (ed. Hedenquist, J. W.) 135–140 (Rep. No. 279, Geol. Surv. Japan, Tsukuba, 1992).
71. Hemley, J. J., Cygan, G. L., Fein, J. B., Robinson, G. R. & d'Angelo, W. M. *Econ. Geol.* **87**, 1–22 (1992).
72. Richards, J. P. *Geology* **20**, 547–550 (1992).
73. van Leeuwen, Th. M., Leach, T., Hawke, A. & Hawke, M. M. *J. Geochim. Explor.* **35**, 1–61 (1990).
74. Torgerson, T. *EOS* **71**, 1, 5, 13 (1990).
75. Crowe, D. E., Valley, J. W. & Baker, K. L. *Geochim. cosmochim. Acta* **54**, 2075–2092 (1990).
76. Heinrich, C. A., Ryan, C. G., Mernagh, T. P. & Eadington, P. J. *Econ. Geol.* **87**, 1566–1583 (1992).
77. Lepel, E. A., Stefansson, K. M. & Zoller, W. H. *J. geophys. Res.* **83**, 6213–6220 (1978).
78. Buat-Menard, P. & Arnold, M. *Geophys. Res. Lett.* **5**, 245–248 (1978).
79. Andres, R. J., Kyle, P. R. & Chuan, R. L. *Geol. Rdsch.* **82**, 687–695 (1993).
80. Krauskopf, K. B. *Introduction to Geochemistry* (McGraw-Hill, New York, 1979).
81. Gilmour, P. in *Advances in Geology of the Porphyry Copper Deposits* (ed. Titley, S. R.) 7–35 (Univ. Arizona Press, Tucson, 1982).
82. Pitzer, K. S. & Pabalan, R. T. *Geochim. cosmochim. Acta* **50**, 1445–1454 (1986).
83. Hedenquist, J. W., Simmons, S. F., Giggenbach, W. F. & Eldridge, C. S. *Geology* **21**, 731–734 (1993).
84. Craig, H. in *Nuclear Geology in Geothermal Area* (ed. Tongiorgi, E.) 17–53 (Consiglio Nazionale delle Ricerche Laboratorio di Geologia Nucleare, Pisa, 1963).
85. Sheppard, S. M. F. in *Stable Isotopes in High Temperature Geological Processes* (eds Valley, J. W., Taylor, H. P. Jr & O'Neil, J. R.) 165–183 (Miner. Soc. Am., Washington DC, 1986).
86. Hedenquist, J. W. in *4th Circum-Pacific Energy and Mineral Resources Conf. Trans.* (ed. Horn, M. K.) 513–524 (Am. Ass. Petrol. Geol., Tulsa, 1987).

ACKNOWLEDGEMENTS. We thank Y. Matsuhisa and members of the Japan-US Seminar on Magmatic Contributions to Hydrothermal Systems for providing the impetus to write this paper; P. Allard and B. W. Christenson allowed use of their unpublished data; E. Lougee helped draft the figures and E. Roedder contributed the photographs; W. F. Giggenbach, R. W. Henley, H. Shinohara and R. H. Sillitoe shared their insight, and A. T. Anderson Jr, A. Arribas Jr, R. O. Fournier, J. J. Rytuba, R. H. Sillitoe, B. E. Taylor and N. C. White provided abundant critical comments on the manuscript.

Phosphatidylinositol-3-OH kinase as a direct target of Ras

**Pablo Rodríguez-Viciana^{*}, Patricia H. Warne^{*}, Ritu Dhand[†],
Bart Vanhaesebroeck[†], Ivan Gout[†], Michael J. Fry[†],
Michael D. Waterfield^{†‡} & Julian Downward^{*§}**

^{*} Imperial Cancer Research Fund, 44 Lincoln's Inn Fields, London WC2A 3PX, UK

[†] Ludwig Institute for Cancer Research, 91 Riding House Street, London W1P 8BT, UK

[‡] Department of Biochemistry and Molecular Biology, University College, Gower Street, London WC1E 6BT, UK

Ras (p21^{ras}) interacts directly with the catalytic subunit of phosphatidylinositol-3-OH kinase in a GTP-dependent manner through the Ras effector site. *In vivo*, dominant negative Ras mutant N17 inhibits growth factor induced production of 3' phosphorylated phosphoinositides in PC12 cells, and transfection of Ras, but not Raf, into COS cells results in a large elevation in the level of these lipids. Therefore Ras can probably regulate phosphatidylinositol-3-OH kinase, providing a point of divergence in signalling pathways downstream of Ras.

THE Ras proteins are key regulators of cell growth and other functions¹; the elucidation of the identities of their critical cellular targets, or 'effectors' through which they exert their biological effects has long been a major goal of research into the regulation of eukaryotic cellular proliferation. The interaction of Ras proteins with their effectors occurs through a stretch of amino acids (32–40), generally known as the effector region, which assumes a different conformation in GTP-bound (active) Ras compared to GDP-bound (inactive) Ras. An effector of Ras would be expected to possess, or control, some sort of enzymatic activity that is regulated by its interaction with Ras.GTP, and which is responsible for at least some of the effects of Ras on the cell. It is possible, indeed likely, that Ras has multiple effectors.

The first protein for which evidence for a possible Ras effector role was found was p12^{GAP} (ref. 2). This protein interacts only with GTP-bound Ras and fails to interact with most effector region mutants. This is also true for the GAP-related protein neurofibromin, the product of the *NF1* gene³. Because both these

proteins strongly stimulate the GTPase activity of Ras they are both negative regulators of Ras: this activity could be incorporated in an effector as a negative feedback mechanism. Some reports indicate that p12^{GAP} can have an influence downstream of Ras in various signalling pathways^{4–7}, but it seems unlikely that it can account for all the signals emerging from Ras. The effectors discovered so far that are most likely to mediate the growth stimulatory signals from Ras are the cytoplasmic serine/threonine Raf kinases. These function downstream of Ras and activate the MAP kinase, and possibly other, pathways⁸. It is not clear whether Raf can activate all the signalling pathways that Ras can, but it can act as a very potent oncogene, unlike GAP. Recently several groups have reported that Ras directly interacts in a GTP-dependent manner with the amino-terminal domain of Raf through its effector region^{9–13}. It remains to be determined whether interaction of Raf with Ras.GTP can result in an increase of its kinase activity.

Raf kinases may not be the sole mediators of the effects of Ras on cellular behaviour. Phosphatidylinositol-3-OH kinase (PI(3)K) coimmunoprecipitates with Ras^{35,36}. This suggests that

§ To whom correspondence should be addressed.

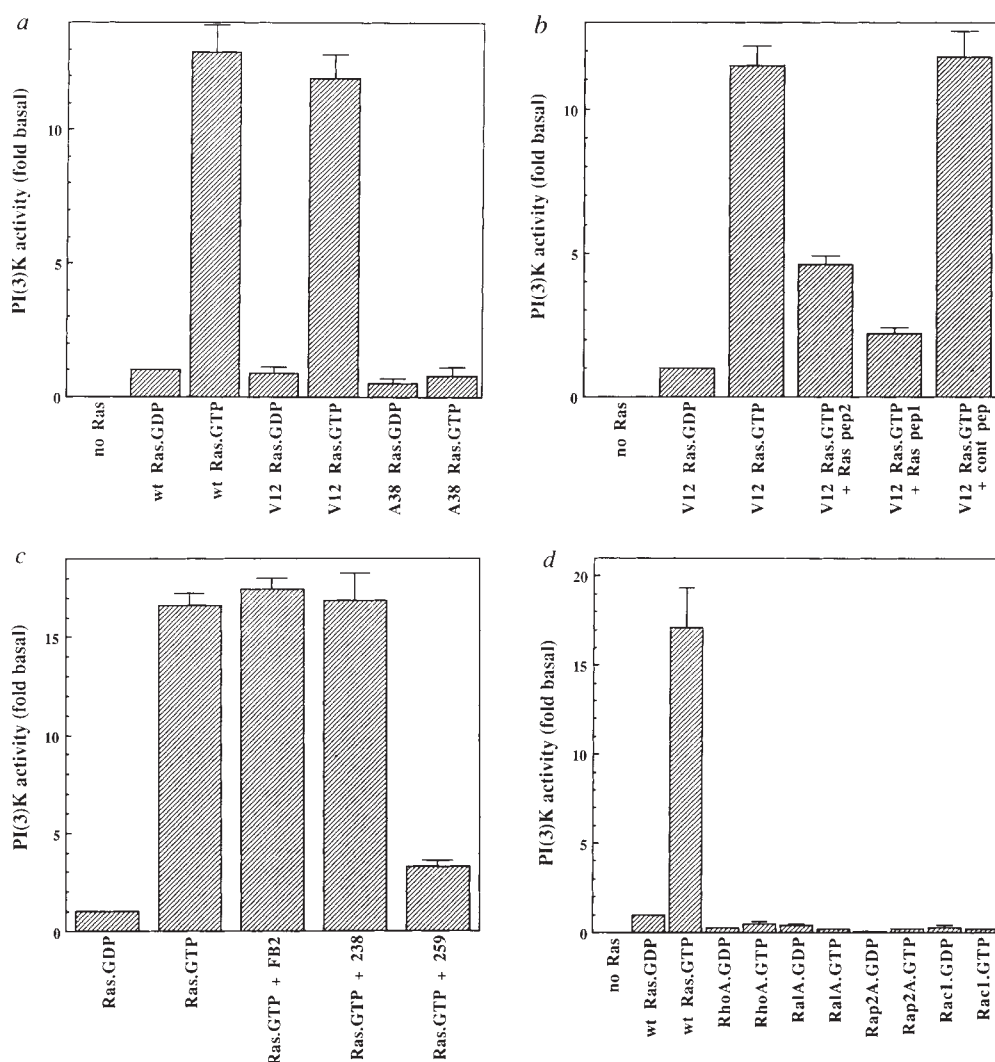
this protein also could be acting either as an effector or as a regulator of Ras. PI(3)K is responsible for phosphorylating phosphoinositides at the 3' position; the generation of 3' phosphorylated inositol phospholipids is induced by many growth factors, including most that activate Ras^{1,21}. Although the role of these lipids is unknown, it seems reasonable that they will play an important part in the regulation of cellular growth and other functions. In this report we present evidence that Ras acts as a direct regulator of PI(3)K.

Interaction of Ras and PI(3)K *in vitro*

Purified Ha-Ras proteins were directly coupled to activated agarose beads. The immobilized Ras proteins were loaded with either GTP or GDP and then incubated with purified PI(3)K (from an Sf9 cell coinfecting with baculoviruses expressing both p85 α and p110 subunits of the kinase). After washing, the amount of PI(3)K activity bound to the Ras beads was measured. Beads carrying wild-type Ras protein loaded with non-hydrolyzable analogues of GTP bound high levels of PI(3)K activity, at least 13 times more than the same beads bound when the Ras had been loaded with GDP and 60 times more than the

background level of activity that was bound to beads without coupled Ras (Fig. 1a). The transforming Ras mutant Val 12 behaved similarly, but an effector mutant which is biologically inactive, Ala 38, was unable to bind any PI(3)K activity. Because this indicated that the interaction between Ras and PI(3)K occurred at the effector region of Ras, the ability of peptides based on this region to inhibit this interaction was studied. A peptide based on Ras residues 17 to 42 (Ras peptide 1) reduced the amount of PI(3)K activity associating with immobilized Ras by about sixfold, whereas a shorter peptide, spanning residues 30 to 42 (Ras peptide 2) reduced the associated activity by about threefold (Fig. 1b). An unrelated peptide did not affect the interaction. These peptides also inhibit other interactions at the Ras effector site such as with p120^{GAP} and Raf^{12,17}. To characterize further the site on Ras that is responsible for this interaction, the immobilized Ras was incubated with various monoclonal antibodies before interaction with PI(3)K. A control antibody (FB2, anti-phosphotyrosine) and a non-neutralizing anti-Ras antibody, Y13-238, have no effect on the amount of PI(3)K binding to Ras, whereas the neutralizing anti-Ras antibody Y13-259¹⁸ greatly inhibits the interaction (Fig. 1c).

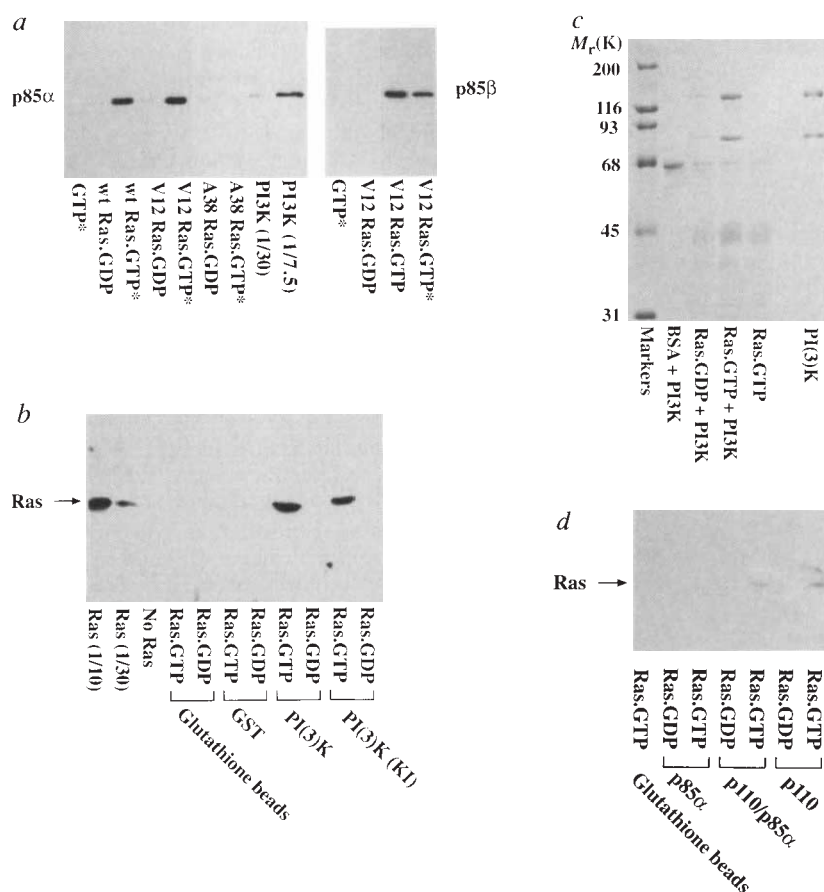
FIG. 1 Association of PI(3)K activity with Ras *in vitro*. a, Immobilized Ras proteins were loaded with GDP or non-hydrolyzable GTP analogue (GMP-PNP) and incubated with PI(3)K. The amount of activity stably associated with the beads is shown. Activity bound to uncoupled beads was subtracted from each value and the corrected activity bound to Ras.GDP beads normalized to 1.0. On this scale the background subtracted was about 0.2. Assays were in duplicate with standard errors shown. b, Inhibition of the association of PI(3)K activity with Ras by peptides derived from the effector site of Ras. Peptides were used at 50 μ M. Peptide 1, H-Ras residues 17–42; peptide 2, H-Ras residues 30–42; control peptide, Sos1 residues 100–120. c, Inhibition of the association of PI(3)K activity with Ras by monoclonal antibodies. d, PI(3)K activity associating with Ras-related proteins. METHODS. Purification of proteins: wild-type Ras and related proteins were purified from baculovirus as described in ref. 30. A38 Ras and RalA were purified from bacteria following the same protocol. PI(3)K was purified from Sf9 cells coinfecting with baculovirus encoding p85 and baculovirus encoding p110. Cell were lysed as in ref. 31. The cytosolic supernatant was loaded into a Superdex-200 column and fractions were analysed by lipid kinase assays on p85 immunoprecipitates. The peak of activity was pooled and used for further experiments. Alternatively (c) cells were infected with a GST–p110 and p85 baculoviruses and the GST–p110/p85 complex as well as other GST-fusion proteins were purified according to the manufacturers protocol. Ras and Ras-related proteins were coupled covalently to Affigel-10 beads (BioRad) following the manufacturers protocol. Interaction of Ras with PI(3)K and lipid kinase assay *in vitro*: purified PI(3)K was added to Ras beads and incubated at 4 °C on a wheel for 2 h.



When doing competition experiments the Ras beads were preincubated with peptides or proteins for 2 h and then PI(3)K was added for a further 2 h. The beads were washed as described in ref. 35 with the inclusion of 5 mM MgCl₂ in all buffers and processed as in ref. 32. Spots on the TLC plate were visualized and quantified using an Ambis scanner.

FIG. 2 Association of p85/p110 with Ras *in vitro*. *a*, Purified p85 α /GST-p110 (left) or p85 β /GST-p110 (right) was incubated with immobilized Ras or mutants of Ras which had been bound to either GDP, GTP or a non-hydrolysable analogue of GTP, GMP-PNP (GTP*). Stably associated PI(3)K was visualized by western blotting using anti-p58 α or anti-p85 β monoclonal antibody. The two right-hand tracks of the left panel show the purified p85 α /p110 preparation used (1/30 and 1/7.5 of the amount used in each binding reaction). *b*, Purified p85 α /GST-p110 was immobilized on glutathione agarose and incubated with Ras which had been bound to either GDP or the non-hydrolysable analogue of GTP, GMP-PNP. Stably associated Ras was visualized by western blotting. The two left-hand tracks show the purified Ras preparation used (1/30 and 1/100 of the amount used in each binding reaction). *c*, Purified p85 α /GST-p110 was incubated with immobilized Ras. Stably associated PI(3)K was visualized by gel electrophoresis followed by staining with Coomassie brilliant blue. The right-hand track shows the purified PI(3)K preparation used (1/10 of the amount used in each binding assay). *d*, Purified GST-p85 α , GST-p110 or GST-p110/p85 α were immobilized on glutathione agarose and incubated with Ras that had been bound to either GDP or GTP. Stably associated Ras was visualized by western blotting.

METHODS. Interaction of Ras and PI(3)K was performed as in Fig. 1. Western blots were done using monoclonal antibody against p85 α , p85 β or Ras using standard protocols³³ with development by enhanced chemiluminescence (Amersham).



The ability of PI(3)K activity to associate with proteins related to Ras was also tested in this assay. Neither of the Rho family proteins tested, RhoA and Rac1, could associate with PI(3)K (Fig. 1*d*). From the Ras family, only Ras itself was positive whereas Ra1A and Rap2A failed to show association.

To confirm the association of PI(3)K with Ras.GTP, the interaction of purified Ras with purified p85/p110 was studied directly by western blotting for the p85 α subunit. In these experiments, p110 was expressed in a baculovirus construct in Sf9 cells as a glutathione *S* transferase fusion protein, along with normal p85. The active p85 α /GST-p110 complex was then purified in a single step on glutathione agarose. p85 α /GST-p110 complexes with wild-type or Val 12 H-Ras when it is bound to GTP, but binds very much more weakly to the GDP bound form (Fig. 2*a*). No interaction is seen between p85 and the effector mutant Ala 38. The proportion of added PI(3)K binding to Ras in these and similar experiments indicates that the affinity of the interaction is in the low micromolar range, similar to that for the interaction of Ras and p120^{GAP}, although the rate of dissociation of the complex is slow. When p85 β /GST-p110 is used in these experiments, it also binds to Ras in a GTP-dependent manner. Similar results are found when p85 α /p110 is used that is not fused to GST (data not shown).

Another way in which the interaction of Ras and PI(3)K can be demonstrated is by incubating soluble Ras, bound either to GTP or GDP, with p85 α /GST-p110 that has been immobilized on glutathione agarose. Bound Ras is visualized by western blotting. Ras does not bind to control or GST agarose beads, but interacts in a GTP-dependent manner with p85 α /GST-p110 beads (Fig. 2*b*). In addition, Ras binds to a point-mutated PI(3)K whose p110 subunit lacks lipid kinase activity (R916P, I. Hiles and M.D.W., unpublished observations), indicating that the enzymatic activity of the kinase is not required for its interaction with Ras.

The above data show that purified Ras interacts with purified PI(3)K in a GTP-dependent manner, but does not rule out the

possibility that this interaction occurs through a contaminating protein. To address this, purified p85 α /GST-p110, which was apparently homogeneous, was bound to purified Ras.GTP immobilized on agarose beads, washed, and the bound proteins analysed by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) followed by staining with Coomassie brilliant blue. The p85 and GST-p110 subunits can be seen binding to the Ras beads (Fig. 2*c*). No other proteins from the PI(3)K preparation are accumulated on the Ras beads, indicating that it is not a subfraction of p85/p110 that is bound to a contaminating protein that is being retained on Ras. The only other proteins present are found in the Ras agarose bead preparation: they are bovine serum albumin (BSA) at 68K, which comes from prewashing the beads with buffers containing BSA to reduce nonspecific binding, and very broad bands below 45 K which are recognized by Ras-specific antibodies and are probably oligomerized Ras fragments (data not shown). It is therefore likely that Ras interacts directly with PI(3)K.

In order to determine whether the interaction of Ras with PI(3)K is through the catalytic p110 subunit or the regulatory p85 subunit, Ras bound either to GTP or GDP was incubated with glutathione agarose immobilized GST-p110, GST-p85 α or GST-p110/p85 α . As shown in Fig. 2*d*, Ras.GTP bound most efficiently to p110 alone, slightly less well to p85 α /p110 complex and not at all to p85 α alone. The interaction of Ras with PI(3)K is therefore mediated through the p110 catalytic subunit.

Ras inhibits elevation of PI(3)K lipid levels

To investigate the possible physiological significance of the ability of GTP-bound Ras to interact with PI(3)K, the effect of expressing a dominant negative mutant of Ras, Asn 17, on growth factor stimulation of phosphatidylinositol 3' phosphorylated lipid levels in intact cells was studied. Because most cell types cannot proliferate while expressing Asn 17 Ras, we chose to use the rat pheochromocytoma cell line PC12 in which Ras regulates differentiation rather than proliferation. PC12 cells

were transfected with a Rous sarcoma virus (RSV) Asn 17 Ras construct and clones of cells selected. The ability of nerve growth factor and epidermal growth factor to elevate phosphatidylinositol 3' phosphorylated lipid levels in these cells was compared to the parental cells. Cells were metabolically labelled with ^{32}P -orthophosphate and levels of the 3' phosphorylated phosphoinositides determined after nerve growth factor (NGF) or epidermal growth factor (EGF) stimulation. The amounts of phosphatidylinositol (3,4) bisphosphate ($\text{PtdIns}(3,4)\text{P}_2$) and phosphatidylinositol (3,4,5) trisphosphate ($\text{PtdIns}(3,4,5)\text{P}_3$) in wild-type PC12 cells increases by between 10- and 20-fold in response to EGF or NGF (Fig. 3a, b). By contrast, in the Asn 17 Ras-expressing PC12 cells, the growth factor stimulation of $\text{PtdIns}(3,4)\text{P}_2$ levels is attenuated by about fivefold. The response of $\text{PtdIns}(3,4,5)\text{P}_3$ lipid pool to growth factors is attenuated somewhat less, by about threefold, because of the expression of Asn 17 Ras in PC12 cells. The incomplete inhibition probably reflects the fact that several mechanisms exist whereby growth factors can control PI(3)K activity, only one of which involves Ras. The time courses of the responses are similar in both wild-type and Asn-17 Ras expressing PC12 cells. Although the simplest explanation for these data, given our *in vitro* observations reported above, is that Ras is involved in the activation PI(3)K, it cannot be ruled out that Ras is acting at the level of lipid breakdown. Active Ras could cause inhibition of phosphoinositide 3' phosphatases.

To be certain that the inhibition of the growth factor induced elevation of $\text{PtdIns}(3,4)\text{P}_2$ and $\text{PtdIns}(3,4,5)\text{P}_3$ in dominant negative Ras-expressing cells is specific to this pathway and not a general phenomenon affecting other signalling pathways, a number of other growth factor-stimulated events were studied. The amount of PI(3)K activity found in anti-phosphotyrosine immunoprecipitates from stimulated cells was measured in an *in vitro* assay. EGF and NGF treatment of PC12 cells lead to a roughly four- and threefold stimulation, respectively, of the amount of PI(3)K activity found in anti-phosphotyrosine immunoprecipitates (Fig. 3c). Expression of Asn 17 Ras in the PC12 cells does not reduce this activity, indicating that growth factor control of PI(3)K through association with phosphotyrosine-containing proteins is normal even in the presence of dominant negative mutant Ras protein. In addition, the production of inositol phosphates (inositol(1,4,5)trisphosphate and inositol(1,4)bisphosphate) by phosphatidylinositol-specific phospholipase C in response to EGF and NGF is also unaltered by Asn 17 Ras expression (see Fig. 3d). Furthermore, expression of Asn 17 Ras in PC12 cells does not obviously alter the ability of EGF and NGF to induce tyrosine phosphorylation of total cellular proteins (see Fig. 3e). It is therefore likely that the expression and function of growth factor receptor tyrosine kinases and the regulation of PI(3)K by tyrosine phosphoprotein binding are not being directly altered by the presence of Asn 17 Ras in these PC12 cells.

Ras elevates phosphorylated lipid levels

From the ability of Ras and PI(3)K to interact *in vitro* and the inhibitory effect of Asn 17 Ras on phosphatidylinositol 3' phosphorylated lipid accumulation *in vivo*, it seems likely that Ras is regulating the function of PI(3)K in intact cells. To explore this possibility further, the effect of overexpressing Ras transiently on 3' phosphorylated phosphoinositides in COS cells was investigated. When a vector expressing activated mutant Val 12 Ras was transfected into COS cells, some elevation was found in the levels of $\text{PtdIns}(3,4)\text{P}_2$ and $\text{PtdIns}(3,4,5)\text{P}_3$ in the intact cells (Fig. 4a). Expression of the p85 α -subunit plus the p110 subunit of PI(3)K together had a similar effect on these lipids. However, when all three proteins were expressed together they cooperated to give a very large increase in the level of 3' phosphorylated phosphoinositides. Wild-type Ras is almost as potent as Val 12 Ras in elevating these lipids (Fig. 4b) presumably reflecting the very high levels of Ras protein expression

achieved in this system, whereas the effector mutant Val 12→Ala 38 Ras is inactive. *v*-Src also caused a similar level of activation. A number of other Ras-related proteins were studied in this assay, including RhoA and Rac1. None of these have any effect on the 3' phosphorylated phosphoinositide levels. The ability of *v*-Src to elevate the levels of these lipids was partially blocked by the dominant negative mutant of Ras, Asn 17, consistent with the hypothesis that *v*-Src can activate PI(3)K through both a Ras-dependent and a Ras-independent pathway, the latter presumably involving tyrosine phosphoprotein binding to p85. The *v*-Raf serine/threonine kinase, or an amino-terminally truncated, activated version of c-Raf (ΔRaf), which act immediately downstream of Ras to activate the MAP kinase pathway, do not cause elevation of 3' phosphorylated phosphoinositides, so it is unlikely that the effects seen with Ras are indirect through secondary events such as autocrine production of growth factors (see also discussion). As also reported previously by others^{19,20}, in this system transfection of either activated Ras or Raf could be shown to activate MAP kinase (data not shown). As in Fig. 3, we cannot at this stage rule out a role for Ras in the control of the rate of breakdown of 3' phosphorylated phosphoinositides, but feel the simplest explanation is that Ras directly contributes to the activation of PI(3)K.

Discussion

The regulation of PI(3)K in response to growth factor treatment of cells could involve a number of mechanisms: p85 is phosphorylated on tyrosine residues in response to treatment of cells with growth factors such as PDGF and it binds through its SH2 domains to autophosphorylation sites on receptor tyrosine kinases and related sites on other molecules, such as IRS-1, CD19 and CD28²¹⁻²³. Tyrosine phosphorylation of p85 has not been shown to change its catalytic activity, but binding to the tyrosine phosphorylated insulin receptor substrate, IRS-1, and to tyrosine phosphorylated peptides, particularly those based on the sequences of the p85 binding sites in the PDGF receptor and IRS-1, does lead to a relatively small increase in the catalytic activity of PI(3)K²⁴⁻²⁶. It is possible that in the intact cell, much stronger stimulations result from such interactions because of translocation of PI(3)K from the cytosol to the plasma membrane where its substrate lipids are located. In addition, another mechanism of regulation could involve proline-rich sequences in p85 binding to SH3 domains of Src-family kinases such as Lck^{14,15}.

Most mitogenic stimuli activate both Ras and PI(3)K^{1,21}. This could reflect the fact that there are common components, such as tyrosine kinases, in the upstream regulatory pathways for both proteins. It could also be the result of a system in which Ras is capable, directly or indirectly, of regulating PI(3)K or *vice versa*. There has been a report that phosphorylation of PDGF receptor at sites involved in the binding of p85 is required for the activation of Ras by PDGF²⁷. However, this appears to occur only in certain cell types²⁸. Furthermore, another adaptor protein, Nck, is now known also to bind to one of these sites²⁹. It therefore seems unlikely that PI(3)K acts upstream of Ras.

The alternative possibility, that Ras acts to control PI(3)K, receives support from the data that are presented here. The association of PI(3)K activity with Ras was previously reported by Lapetina³⁵, who also showed that p85 could be detected in association with Ras by immunoblotting³⁶. Here we have extended these observations to show that the interaction of Ras and PI(3)K is direct, requires Ras to be in the active, GTP-bound conformation and occurs through the effector site on Ras. Ras can therefore interact with PI(3)K *in vitro* in the manner expected for a true effector interaction. We have not yet been able to reproducibly demonstrate that this interaction results in an increase in the lipid kinase activity. The original observation of phosphatidylinositol activity in Y13-259 immunoprecipitates of Ras³⁵ indicated that it was unlikely that the interaction was through the effector site on Ras because this antibody blocks

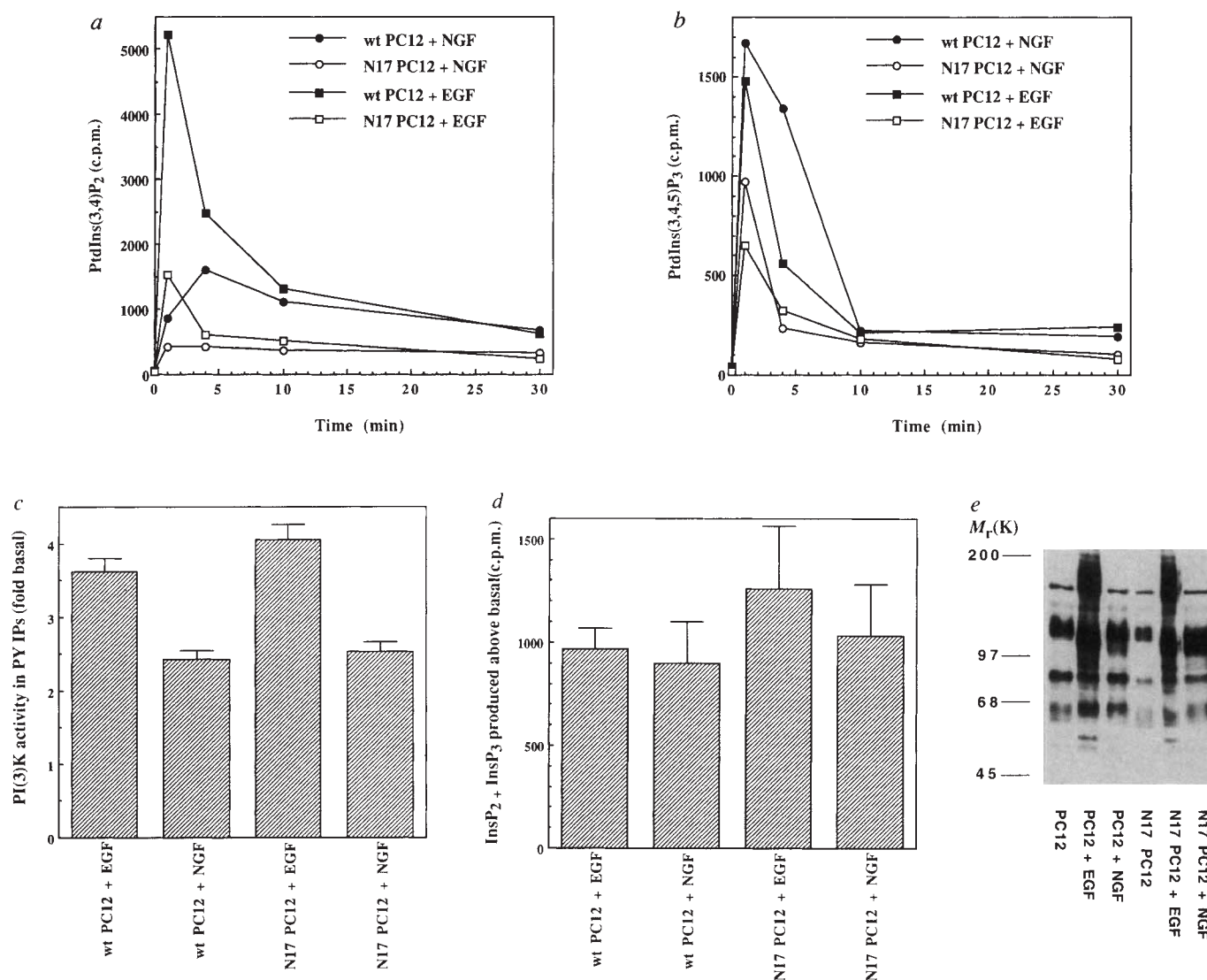


FIG. 3 Dominant negative Ras inhibition of growth factor stimulation of 3' phosphorylated phosphoinositide levels in intact PC12 cells. **a**, Wild-type PC12 and Asn 17 Ras-expressing PC12 cells were metabolically labelled with ³²P-orthophosphate before stimulation with EGF or NGF. Phospholipids were extracted from the cells and the level of PtdIns(3,4)P₂ determined. Representative of four independent experiments. **b**, As for **a**, except that levels of PtdIns(3,4,5)P₃ are shown. **c**, Anti-phosphotyrosine immunoprecipitates (PY 1Ps) were made from PC12 cells and Asn 17 Ras-expressing PC12 cells treated with EGF or NGF. These immunoprecipitates were assayed for PI(3)K activity *in vitro*. Activity is expressed as multiples of the amount in similar immunoprecipitates from unstimulated cells. **d**, Wild-type PC12 and Asn 17 Ras-expressing PC12 cells were metabolically labelled with ³H-inositol before stimulation with EGF for 1 min or NGF for 4 min. Soluble inositol phosphates were extracted from the cells and the levels of inositol(1,4) bisphosphate plus inositol(1,4,5)trisphosphate were determined. The elevation of inositol polyphosphate levels above those in unstimulated cells is shown. Representative of four independent experiments. **e**, Anti-phosphotyrosine western blots of whole cell lysates from PC12 cells and Asn 17 PC12 cells treated with NGF (for 4 min) or EGF (for 1 min).

METHODS. Generation of PC12 clones expressing N17 H-Ras: PC12 cells were transfected by the calcium phosphate method with RSVneoN17 H-Ras and G418 resistant clones were picked and analysed by western blot. Two clones, 7 and 8, expressing highest levels of N17-Ras were chosen for further experiments. All were insensitive to NGF-induced differentiation. *In vivo* measurement of PI(3)K activity: 60-mm dishes of subconfluent PC12 cells were labelled with ³²P-orthophosphate (250 μ Ci per dish) and PI(3)K products analysed³⁴. The dried combined CHCl₃ phase was normalized for radioactivity, redissolved and run on an oxalated silica gel 60 TLC plate, using a CHCl₃:methanol:acetone:acetic acid:water (40:13:15:12:7) system. The area corresponding to PtdIns(3,4)P₂ and PtdIns(3,4,5)P₃ was scraped off and deacylated by incubating for 60 min at 50 °C with methylamine. Deacylated products were re-extracted with water, subjected to HPLC analysis and fractions counted by Cerenkov emission. Western blots: PC12 cells were lysed in 1% Triton X-100, 25 mM Tris 7.5, 100 mM NaCl, 1 mM EDTA, 1 mM EGTA, 20 mM NaF, 1 mM Na₃VO₄, 1 mM PNPP, 1 mM PMSF, 10 μ g ml⁻¹ aprotinin, 10 μ g ml⁻¹ leupeptin. Western blots were done using 4G10 anti-phosphotyrosine monoclonal antibody following the manufacturer's protocol (UBI).

its biological activity. However, we find that Y13-259 strongly inhibits the interaction of Ras and PI(3)K: it is possible that differences exist between various preparations of this antibody, or that it does not completely inhibit the interaction.

The question of whether Ras is involved in the regulation of PI(3)K within intact cells is critical. The demonstration here that a dominant negative Ras mutant inhibits the ability of NGF

and EGF to elevate phosphatidylinositol 3' phosphorylated lipid levels in PC12 cells *in vivo* suggests that Ras acts upstream of the lipid kinase. Further *in vivo* evidence implicating Ras in the regulation of PI(3)K comes from the transient transfections in COS cells: the elevation of cellular 3' phosphorylated phosphoinositides caused by Ras when coexpressed with p85 and p110 indicates that Ras can activate PI(3)K. It is possible that expres-

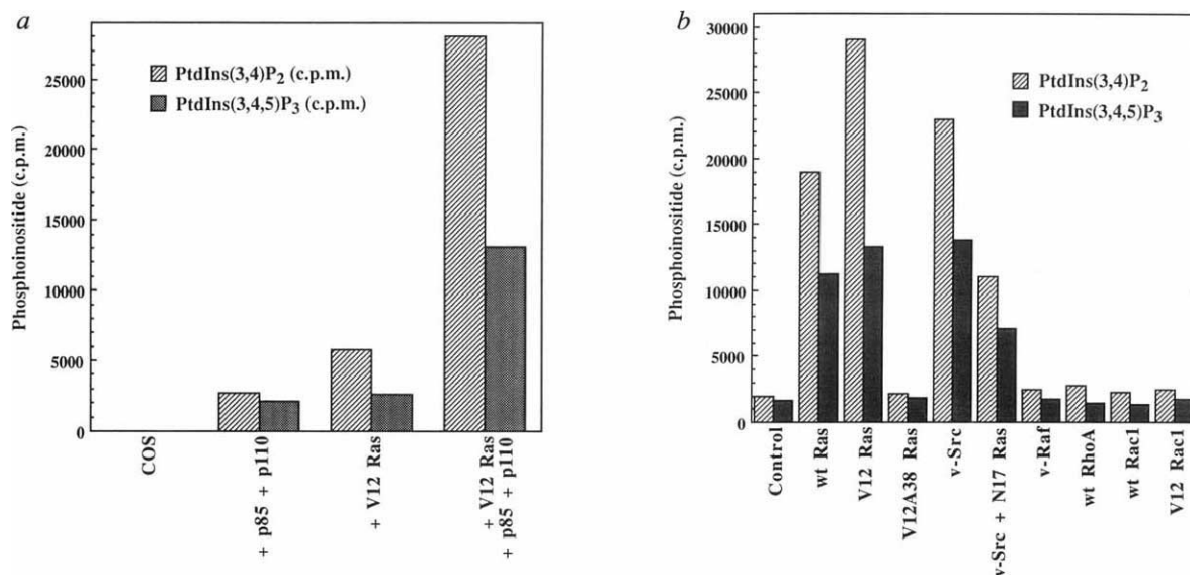


FIG. 4 Elevation of 3' phosphorylated phosphoinositide levels in intact COS cells by expression of Ras with p85 and p110. **a**, COS cells were transfected with the indicated plasmids before ³²P-orthophosphate metabolic labelling and determination of levels of PtdIns(3,4)P₂ and PtdIns(3,4,5)P₃. The level of these lipids in untransfected COS cells is taken as zero. **b**, As in **a** except that all cells have received p85 plus p110 plasmids in addition to the constructs indicated. The first pair of values are for COS cells transfected with p85 and p110 alone.

METHODS. COS-1 were collected, washed in PBS and 6×10^6 cells

were transfected by electroporation (450 V, 250 μ F) with 10 μ g each of pMT2p85 α , pSG5p110 plus all the other plasmids where appropriate. Cells were allowed to stand on ice for 10 min and 1/10 were seeded in a 35 mm dish to check expression of the different constructs by western blot. The rest were seeded in a 10 cm dish for ³²P labelling, 48 h after transfection cells were washed with phosphate-free DMEM and labelled for 4–5 h with 300 μ Ci ³²P-orthophosphate per 10-cm dish. Reactions were terminated and samples processed as above. Expression of each protein was checked by western blotting of whole-cell lysates with the appropriate antibodies.

sion of activated Ras could lead to the autocrine production of growth factors that could cause activation of PI(3)K through a tyrosine kinase mediated pathway. Although the experiments have been designed to minimize the impact of autocrine growth factor production, a more compelling argument for a direct effect of Ras is that activated Raf does not have the same stimulatory effect as Ras on 3' phosphorylated phosphoinositide levels. This is despite the fact that Raf causes production of a similar set of autocrine growth factors as does Ras, presumably by acting as a Ras effector. In the COS cell system used here we find that transfection of either activated Ras or Raf leads to the production of autocrine factors in the medium that are capable of acutely activating microtubule-associated protein (MAP) kinase; however, these factors are not capable of elevating PtdIns(3,4,5)P₃ levels (P.R.-V. and J.D., unpublished observations). The data from the COS cells indicate that there exists a

point of divergence in the signalling pathway downstream of Ras and upstream of the Raf kinase domain which could be accounted for by a direct activation of PI(3)K by Ras.

The data presented here provide evidence for regulation of PI(3)K by Ras both *in vitro* and *in vivo*. A number of mechanisms for regulating PI(3)K therefore exist. Such redundancy in signalling pathways is becoming a common feature in many systems. It is likely that these pathways synergize to cause strong activation of the lipid kinase; the Ras interaction by itself may not be sufficient to activate PI(3)K, but may require p85/p110 to be correctly localized to the plasma membrane through SH2 or SH3 interactions. The importance of the products of PI(3)K in regulation of cell growth is still unknown, so the relevance of the ability of Ras to control it is not yet clear. Further knowledge of the targets of PtdIns(3,4,5)P₃ will be needed before it is possible to estimate its contribution to the effects of Ras on the cell. □

Received 31 March; accepted 21 June 1994.

- Lowy, D. R. & Willumsen, B. M. A. *Rev. Biochem.* **62**, 851–891 (1993).
- Trahey, M. & McCormick, F. *Science* **238**, 542–545 (1987).
- Downward, J. *Curr. Biol.* **1**, 353–355 (1991).
- Duchesne, M. et al. *Science* **259**, 525–528 (1993).
- Medema, R. H. et al. *Molec. cell. Biol.* **12**, 3425–3430 (1992).
- Schweighoffer, F. et al. *Science* **256**, 825–827 (1992).
- Yatani, A. et al. *Cell* **61**, 769–776 (1990).
- Marshall, C. J. *Curr. Opin. Genet. Dev.* **4**, 82–89 (1994).
- Zhang, X.-F. et al. *Nature* **364**, 308–313 (1993).
- Vojtek, A. B., Hollenberg, S. M. & Cooper, J. A. *Cell* **74**, 205–214 (1993).
- Van Aelst, L., Barr, M., Marcus, S., Polverino, A. & Wigler, M. *Proc. natn. Acad. Sci. U.S.A.* **90**, 6213–6217 (1993).
- Warne, P. H., Vicinina, P. R. & Downward, J. *Nature* **364**, 352–355 (1993).
- Moodie, S. A., Willumsen, B. M., Weber, M. J. & Wolfman, A. *Science* **260**, 1658–1661 (1993).
- Vogel, L. B. & Fujita, D. J. *Molec. cell. Biol.* **13**, 7408–7417 (1993).
- Prasad, K. V. S. et al. *Molec. cell. Biol.* **13**, 7708–7717 (1993).
- Feramisco, J. R., Gross, M., Kamata, T., Rosenberg, M. & Sweet, R. W. *Cell* **38**, 109–117 (1984).
- Schaber, M. D. et al. *Prot. Struct. Funct. Genet.* **6**, 306–315 (1989).
- Furth, M. E., Davis, L. J., Fleurdelys, B. & Scolnick, E. M. *J. Virol.* **43**, 294–304 (1982).
- Howe, L. R. et al. *Cell* **71**, 335–342 (1992).
- Kyriakis, J. M. et al. *Nature* **358**, 417–421 (1992).

- Stephens, L. R., Jackson, T. R. & Hawkins, P. T. *Biochim. biophys. Acta* **1179**, 27–75 (1993).
- Tuveson, D. A., Carter, R. H., Soltoff, S. P. & Fearon, D. T. *Science* **260**, 986–989 (1993).
- Pages, F. et al. *Nature* **369**, 327–329 (1994).
- Carpenter, C. L. et al. *J. Biol. Chem.* **268**, 9478–9483 (1993).
- Backer, J. M. et al. *EMBO J.* **11**, 3469–3479 (1992).
- Myers, M. G. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 10350–10354 (1992).
- Fanti, W. J. et al. *Cell* **69**, 413–423 (1992).
- Satoh, T. et al. *Molec. cell. Biol.* **13**, 3706–3713 (1993).
- Nishimura, R. et al. *Molec. cell. Biol.* **13**, 6889–6896 (1993).
- Page, M. J. et al. *J. Biol. Chem.* **264**, 19147–19154 (1989).
- Gout, I. et al. *Biochem. J.* **288**, 395–405 (1992).
- Reif, K., Gout, I., Waterfield, M. D. & Cantrell, D. A. *J. Biol. Chem.* **268**, 10780–10788 (1993).
- Harlow, E. & Lane, D. *Antibodies: a Laboratory Manual* (Cold Spring Harbor Laboratory Press, New York, 1988).
- Chung, J., Kuo, C. J., Crabtree, G. R. & Blenis, J. *Cell* **69**, 1227–1236 (1992).
- Sjölinder, A., Yamamoto, K., Huber, B. E. & Lapetina, E. C. *Proc. natn. Acad. Sci. U.S.A.* **88**, 7908–7912 (1991).
- Sjölinder, A. & Lapetina, E. C. *Biochem. biophys. Res. Commun.* **189**, 1503–1508 (1992).

ACKNOWLEDGEMENTS. We thank C. Calés for her support, R. Wocholski and D. Pappin for help with HPLC analysis, C. Allbright and A. Hall for Ras-related protein expression constructs, C. Marshall for Raf constructs and E. Lapetina and P. Parker for helpful discussions. B.V. is supported by a fellowship from the Belgian NFWO.

Structure of ribonucleotide reductase protein R1

Ulla Uhlin & Hans Eklund

Department of Molecular Biology, Swedish University of Agricultural Sciences, Uppsala Biomedical Center, Box 590, S-751 24 Uppsala, Sweden

Ribonucleotide reductase is the only enzyme that catalyses *de novo* formation of deoxyribonucleotides and is thus a key enzyme in DNA synthesis. The radical-based reaction involves five cysteines. Two redox-active cysteines are located at adjacent antiparallel strands in a new type of ten-stranded α/β -barrel, and two others at the carboxyl end in a flexible arm. The fifth cysteine, in a loop in the centre of the barrel, is positioned to initiate the radical reaction.

RIBONUCLEOTIDE reductase (RNR) is an essential enzyme in DNA synthesis because it catalyses all *de novo* synthesis of deoxyribonucleotides. Its activity is cell-cycle dependent and is highly transcriptionally regulated^{1–3}. The overall enzymatic activity is regulated by ATP (activation) and dATP (feedback inhibition at high concentrations, except for viruses)⁴. RNR has a unique additional allosteric regulation because triphosphate nucleotides regulate the substrate specificity in such a way that a balanced supply of the different deoxynucleotides are present during DNA synthesis.

The reduction of a ribonucleotide to a deoxyribonucleotide is a chemically difficult reaction and has to be initiated by an organic free radical^{4,5}. Surprisingly this is done in different ways in different ribonucleotide reductases and at least three different classes of reductases have been characterized⁴. The class II reductases use adenosyl cobalamin, whereas class III anaerobic reductase probably uses a glycine radical generated with the help of an *S*-adenosyl methionine and an iron sulphur centre. However, most of the reductases so far characterized use an iron-containing protein for the generation of the catalytically necessary organic radical (class I). This is true for all mammalian reductases as well as for those from *Escherichia coli* and DNA viruses. In these class I reductases, the molecular organization is an $\alpha_2\beta_2$ holoenzyme where the two homodimers are denoted R1 and R2. In *E. coli*, the R1 protein has a relative molecular mass (M_r) of 171K (2×761 residues) and R2 has a M_r of 87K (2×375 residues). The R1 protein contains the allosteric binding sites and the active site, whereas each subunit of the R2 protein contains a radical located on a tyrosine side chain close to a binuclear iron centre. The three-dimensional structure of the *E. coli* R2 protein has earlier been determined⁶ and refined at 2.2 Å resolution⁷.

However, organic radicals are highly reactive and cannot be allowed freely in cells. The radical reaction must therefore not be initiated until the substrate is bound at the active site of the R1 protein. The tyrosyl radical is buried about 10 Å inside the R2 protein^{6,7} and is stable *in vitro* for weeks at 4 °C. During the enzymatic reaction, the tyrosyl radical has to relocate from its buried position inside the R2 protein, to the active site of the R1 protein. In studies where the substrate was modified⁸ or the active site residues changed by site-directed mutagenesis, the substrate can be decomposed and react with the enzyme, and in some cases cleave the peptide chain^{9–11}. This suggests that the radical transfer has to be localized to a specific pathway⁷.

The substrate is reduced by one pair of cysteines at the active site of R1 which are oxidized during the reaction. Oxidized R1 can be reduced by at least two systems, the thioredoxin and glutaredoxin systems, both of which use NADPH as ultimate reductant¹². A third cysteine (Cys 439 in *E. coli*), essential for the activity, has been suggested to form a radical and abstract

the 3'-hydrogen from the ribose ring of the substrate to create a substrate radical^{9–11}.

The bacterial R2 is a homologue of the mammalian and herpes virus enzymes which are potential drug targets. The carboxyl end peptides of R2 are the basis for drug development¹³ because the formation of the holoenzyme complex can be inhibited by carboxy-terminal peptides from the R2 protein^{14–16}. A synthetic 20-residue peptide corresponding to the C terminus of *E. coli* R2 was used to obtain crystals of R1 suitable for crystallographic investigations¹⁷. In contrast to crystal forms obtained in the absence of peptide, these crystals can be obtained reproducibly, are stable and diffract to 2.5 Å in the synchrotron beam. We report here the structure of the R1 protein in complex with the R2 peptide to 2.5 Å resolution.

Subunit structure

Heavy-atom derivative phases were improved to 3.1 Å resolution by solvent flattening, histogram matching, Sayre's equation¹⁸ and threefold averaging¹⁹ (Table 1) and extended from 3.1 to 2.7 Å resolution. The resulting electron density map was of good quality and the chain could be traced from residue 10 to residue 737. The first 9 and last 24 residues were not visible in the map and three loops near the subunit interaction area had only weak density. The structure has been partly refined to an *R*-factor of 21% at 2.5 Å resolution. Part of the electron density map is shown in Fig. 1 and details of the structure determination are given in Table 1.

The R1 subunit is composed of three domains: one mainly helical N-terminal domain (~220 residues), one α/β barrel domain (~480 residues), and a 70-residue $\alpha\beta\alpha\alpha\beta$ domain (Fig. 2). The active site is located in the centre of a deep cleft across the subunit between the N-terminal and barrel domains.

The barrel domain is of a new type where the central pleated sheet has 10 β -strands. Compared to the common 8-stranded α/β barrels^{20,21}, this barrel is larger and has the β -strands differently arranged. The barrel in R1 is composed of two halves, where each half contains five strands and four regularly arranged connecting helices (Fig. 3). The hydrogen bonding within each half is parallel, whereas the first and last strand of each half-barrel have antiparallel hydrogen bonding to each other (βA to βF and βE to βJ). Thus, the active site has a different location relative to the strands than in the 8-stranded α/β barrels, where the active site is always found at the carboxyl ends of the barrel strands. In the R1 barrel, the strands in the first half of the barrel (βA – βE) have their amino ends towards the active site, whereas the strands of the second half of the barrel (βF – βJ) point in the opposite direction.

The wider R1 barrel allows the insertion of a loop like a finger in its centre. One of the conserved cysteines, 439, sits on the tip of this finger along with the conserved Asn 437, Leu 438 and

Glu 441. This 20-residue loop, which connects the two half-barrels, starts after β E, folds into the middle of the barrel all the way to the top, makes a β -turn, goes back through the barrel and continues across the bottom of the barrel to β F.

The active site

In contrast to most other known thiol redox proteins, the redox-active cysteines of R1 are not located sequentially close to each other (that is, with two or four intervening residues) but are separated by hundreds of residues. The cysteines that have been proposed to be directly involved in the substrate reduction, 225 and 462^{9-11,22-25} are instead held close by their location on two adjacent β -strands in the barrel (β A and β F). They are located at equivalent positions on the first strand of each half-barrel. The third cysteine of the active site, Cys 439, has a central position at the barrel focus, enthroned on the internal loop finger (Fig. 3a). It is thought to form a thiyl radical and be directly involved in the abstraction of the 3' hydrogen from the ribose ring of the substrate^{9,26}. These three active-site cysteines form a triad on one side of the barrel, in the pocket between the N-terminal and barrel domains.

In our crystals the redox-active cysteines (225, 462) are oxidized and form a disulphide bridge. Cys 225 is more accessible to solvent than Cys 462 and also closer to Cys 439. A distance of 6 Å separates the sulphurs of Cys 225 and Cys 439, but they may be brought into van der Waals contact by side-chain torsions. However, Cys 225 and Cys 439 are not close enough to form a disulphide bond and the structure is such that it prevents closer approach of the cysteines.

Most of the conserved residues in R1 are lining the active-site cavity between the two domains. These residues are hydrophobic (Pro 210, Leu 438, Leu 464, Met 620, Pro 621) and polar (Ser 224, Asn 437). No aromatic residues are located in the substrate cavity. The only acidic or basic residue at the active centre is the conserved Glu 441 which may bind the ribose moiety of the nucleotide substrate or act as a base in the enzyme reaction. All three active-site cysteines are about 5 Å from the carboxylate of Glu 441 (Fig. 4a). Asn 437 has a key position in the centre of the active site and appears to form a hydrogen bond to the sulphur of Cys 225 with one of its nitrogen hydrogen atoms, and to Glu 441 with the other. The Asn 437 side chain and the inter-

TABLE 1 Experimental data

(a) X-ray datasets and heavy atom refinement									
Data-set*	Resol. Å	Compl. %	No. unique reflections	R_{sym} (%)	D -iso† (%)	Soak‡ (mM h)	No. of sites	F_H/ϵ § (ace/ce)	R_C § (ace/ce)
Native	2.5	94	109,500	9.8	—	—	—	—	—
KAu(CN) ₂	5.5	88	10,374	6.7	11.3	0.1/24	12	1.1/0.8	0.8/0.8
K ₂ PtCl ₄	5.5	81	9,558	7.8	25.2	3/18	15	1.2/0.8	0.8/0.8
K ₂ Pt(CN) ₄	3.0	59	39,372	3.5	17.5	10/26	12	0.7/0.7	0.9/0.8
Ta ₆ Br ₁₄	3.0	45	30,298	7.2	20.0	0.5/20	6	1.4/1.0	0.7/0.8
Ta ₆ Br ₁₄	3.0	42	26,048	4.9	20.0	0.5/20	6	1.3/0.8	0.8/0.9
UO ₂ Ac ₂	3.0	44	29,307	4.7	19.6	5/95	10	0.7/0.5	0.9/0.9
UO ₂ Ac ₂	3.0	50	33,500	4.2	13.7	5/94	10	0.7/0.6	0.9/0.9
(b) Phase improvement									
Phasing	Resol. (Å)	Figure of merit	R -factor (%)	Correlation (F_o , F_c)					
MIR (Heavy)	15–3.1	0.42	48	0.56					
MIR (Mlphare)	15–3.1	0.36							
Solvent flattening	15–3.1	0.84	42	0.62					
SQUASH	15–3.1	0.91							
Averaging (A)*	15–2.7		22	0.93					
Averaging (B)#	15–2.7		20	0.93					
Averaging (C)**	15–2.7		18	0.95					

To grow good crystals, a modified rapid procedure was essential as well as the inclusion of a synthetic peptide in the crystallization drops. This 20 amino-acid peptide has the same sequence as the C terminus of R2 (YLVGQ IDSEV DTDDL SNFQL (single-letter code)). These crystals of the space group $R32$ with hexagonal cell axes $a=b=226$ Å and $c=341$ Å with three subunits in the asymmetric unit diffract to about 2.5 Å in synchrotron radiation¹⁷. 3.0 Å data were collected on a Nicolet detector using a Rigaku rotating anode source and processed with the Buddha program³⁵. Native data sets to higher resolution collected at synchrotron stations in Hamburg and Daresbury were processed^{36,37} and scaled (CCP4 program suite, Daresbury, England). Difference Pattersons were solved with help of the RSPS program³⁸. Heavy-atom parameters for four derivatives were refined^{39,40} and the phases were further improved by solvent flattening⁴¹, histogram matching and Sayre's equation¹⁸ and then cycled with phase combination in each cycle. A skeletonized representation of the solvent flattened map was calculated and an envelope was defined from the bones of one subunit using the program O (refs 19, 42, 43). Cyclic 3-fold averaging was done using this envelope and the matrices obtained from the heavy-atom parameters^{19,44}. The non-crystallographic operators were optimized (G. J. Kleywegt and T. A. Jones, Uppsala), and the envelope was slightly extended. Averaging was done and the phases extended in steps of 0.05 Å from 3.1 to 2.7 Å resolution. The resulting map was of such quality that most of the molecule could be fitted to the density. Further refinement of the model to 2.5 Å resolution was made in X-plor⁴⁵. At present the R -factor is 0.21 at 2.5 Å with an r.m.s. difference of 0.019 Å from standard bond lengths. 234 well ordered water molecules have been added.

* Native (204,000 observed reflections)=merged data collected on image plate systems at synchrotron station X31 in Hamburg and Daresbury station 9.6. Derivatives collected on a Xenotronics area detector mounted on a Rigaku rotating anode. The main sites of the gold and platinum derivatives are situated close to the cysteines at the active site.

† Scaling against native Xenotronics data in R_{stats} .

‡ The crystals were soaked in the mother liquor, where the heavy metal solution was added to the appropriate concentration.

§ Statistics from Mlphare (tantal derivatives only to 5.0 Å). F_H/ϵ (phasing power)=r.m.s. (F_H)/r.m.s. (lack of closure), where F_H is the calculated heavy-atom structure factor. $R_C = \{ \sum |F_H(\text{obs})| - |F_H(\text{calc})| \} / \sum F_H(\text{obs})$. ace, acentric; ce, centric reflections.

|| To get R_{stats} statistics comparable after different refinement methods, the R -factor was calculated on an average of the three molecules using the same envelope and operators as in average (C).

*|| MIR phases from heavy, solvent flattening (Wang) and 3-fold averaging.

MIR phases from heavy, solvent flattening, histogram matching, Sayre's equation (Squash) and 3-fold averaging.

** Phases from built model were used to locate minor sites and for initial heavy-atom refinement in Mlphare. Squash and 3-fold averaging as in (B).

FIG. 1 Stereo view electron density map ($2|F_o| - |F_c|$ map contoured at 1σ) of the conserved residues (Cys 439 (centre), Tyr 730 and Tyr 731 which we propose are involved in electron transfer from R2 to the substrate. One water molecule is hydrogen bonded to the side chain OH of Tyr 730.

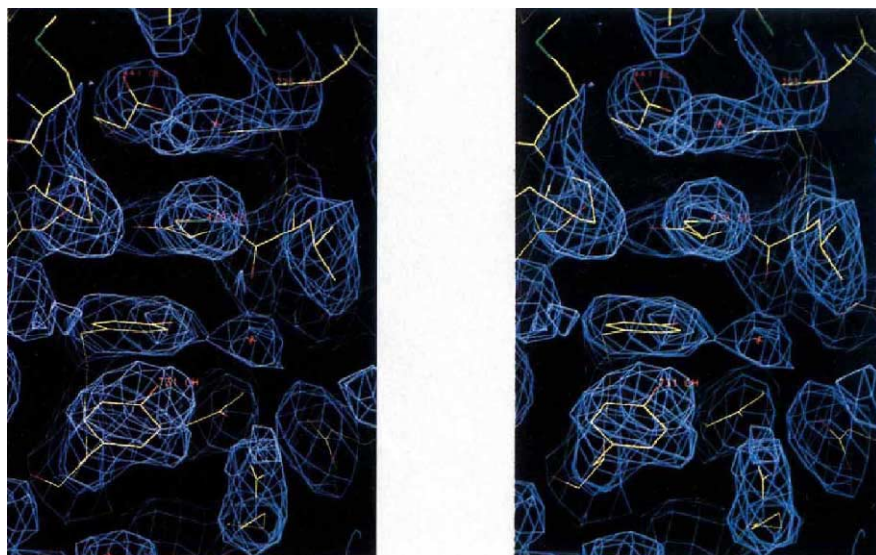


FIG. 2 a, One R1 subunit can be divided into two large domains: the N-terminal domain in blue and the α/β barrel in yellow with the β strands in orange. The small domain in grey is an excursion in the regular barrel after βD . It covers the bottom of the barrel. The N-terminal residues (1–19) and loop 1 have weak density and are not included in the picture. The figures of the molecule were made using Molscript⁴⁶. The R1 subunit viewed from the side can be compared with a side view of a left hand with the fingers at a right angle from the palm. The N-terminus would then be at the tip of the longest finger, and the thumb would illustrate the β -hairpin ($\beta 10$ – $\beta 11$) packed against the 4-helix bundle in the N-terminal domain. The longest helix in the barrel (αG , 28 residues long) is situated as the knuckles between the fingers and the palm. The centre of the palm would correspond to the location of the active site. The N-terminal domain is α -helical (12 α -helices) except for a β -hairpin and a short β -strand connected to the first strand of the second half of the α/β barrel. The first helices of the N-terminal domain form a left-handed 4-helix bundle without overhead connections⁴⁷ very similar to the 4-helix bundle in thermolysin. b, A schematic drawing of the secondary structure elements of the R1 subunit. The regular parts of the barrel are labelled with letters, others are labelled with numbers in consecutive order. 3_{10} helices are also labelled α . L1–L3 are loops 1–3. The location of the three conserved cysteines at the active site are shown.

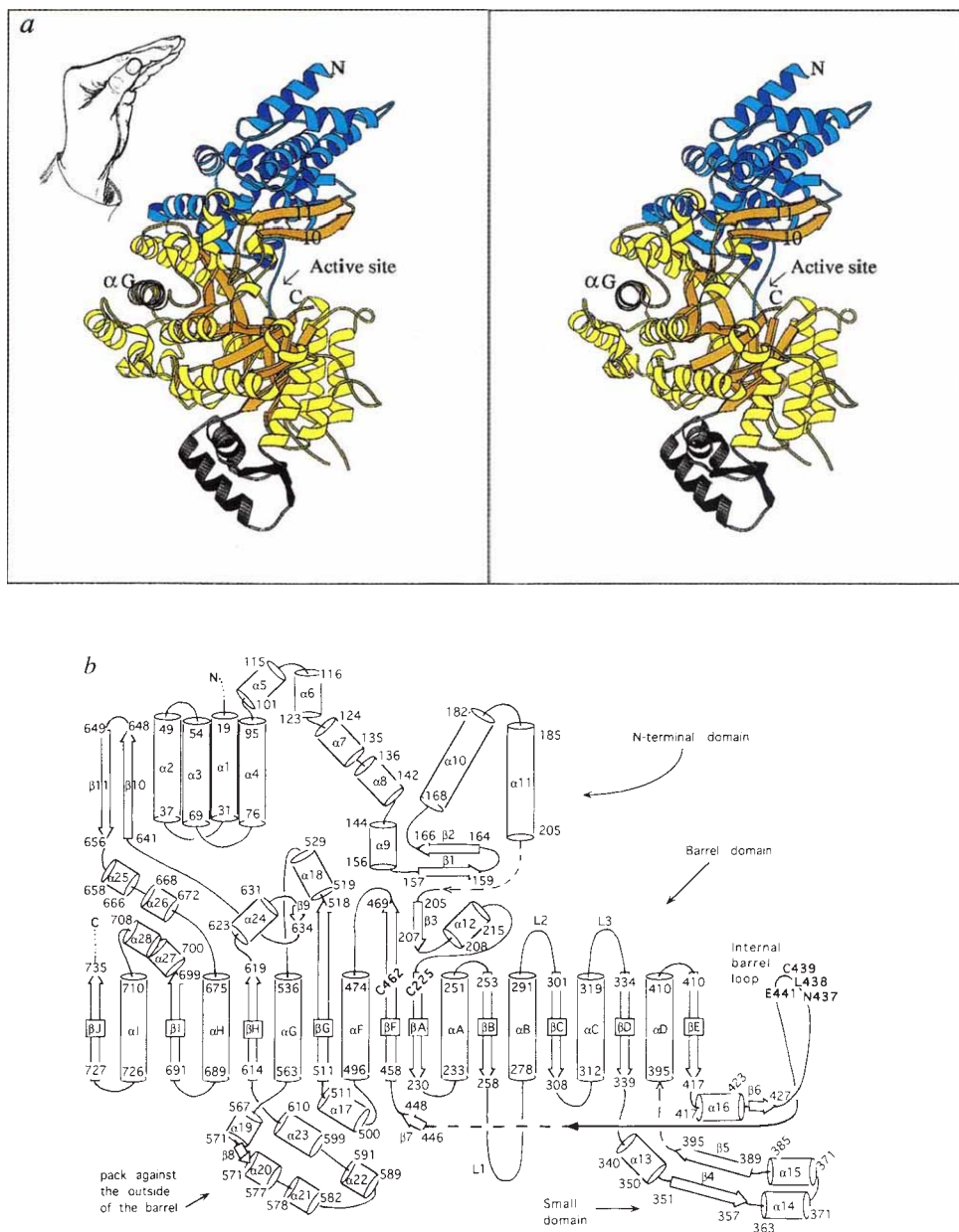


FIG. 3 *a*, A simplified drawing showing the organization of the α/β barrel. The regular elements $\alpha A-\alpha D$ and $\beta A-\beta E$ belong to the first half-barrel, and $\alpha F-\alpha I$ and $\beta F-\beta J$ to the second half-barrel. The two halves are connected by a loop in the centre of the barrel (dashed lines). The two halves are pseudo 2-fold related, having similar fold but different additional intervening secondary structure elements and loops. The double arrow indicates the pseudo 2-fold symmetry of the barrel. This barrel may have evolved from a gene duplication of the two subdomains and gone through an extensive divergence over a long period of time. *b*, The α/β barrel viewed along its β -strands. For clarity, the N-terminal domain, the small domain covering the bottom of the barrel and the long anti-parallel hairpin ($\beta 10-\beta 11$) have been removed. The first half of the barrel is in yellow, the second half in green and the connecting loop in magenta. Extra secondary elements in the second half of the barrel are coloured light green. The β -barrel is elliptical with the dimensions $15 \times 20 \text{ \AA}$. The three conserved cysteines at the active site are labelled. *c*, The binding site for the carboxyl end of the R2-peptide bound to R1. The peptide is labelled starting from the carboxyl end: L1 to D8. The peptide is bound between helices $\alpha 13$ in the small domain and $\alpha 1$ of the barrel. The carboxyl end group of the peptide binds in a positively charged part and forms a hydrogen bond to the side chain of Lys 584. The side chain of the last residue of the peptide sits in a hydrophobic pocket. Lys 723 forms a hydrogen bond to the carbonyl oxygen between the two last residues.

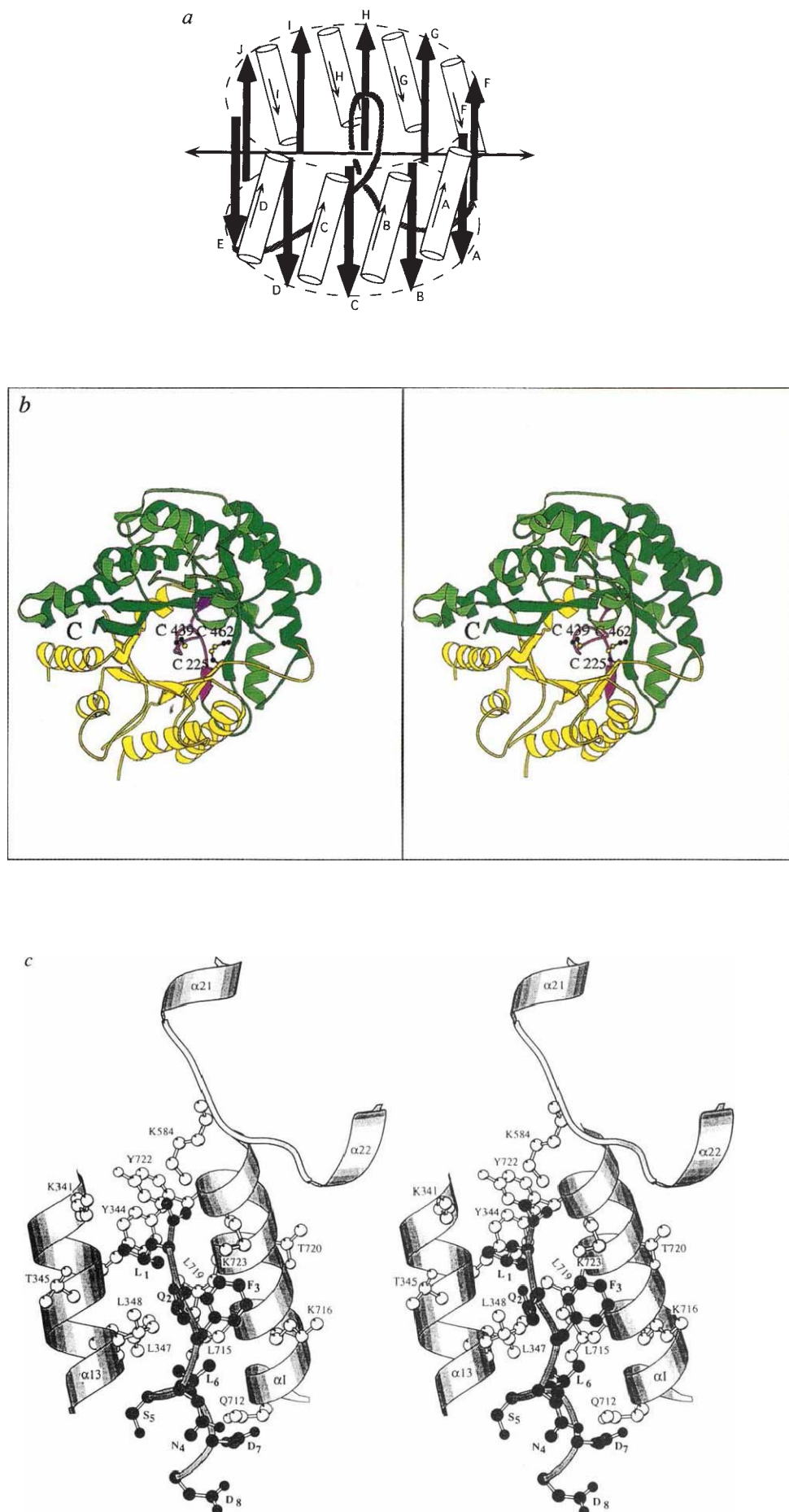
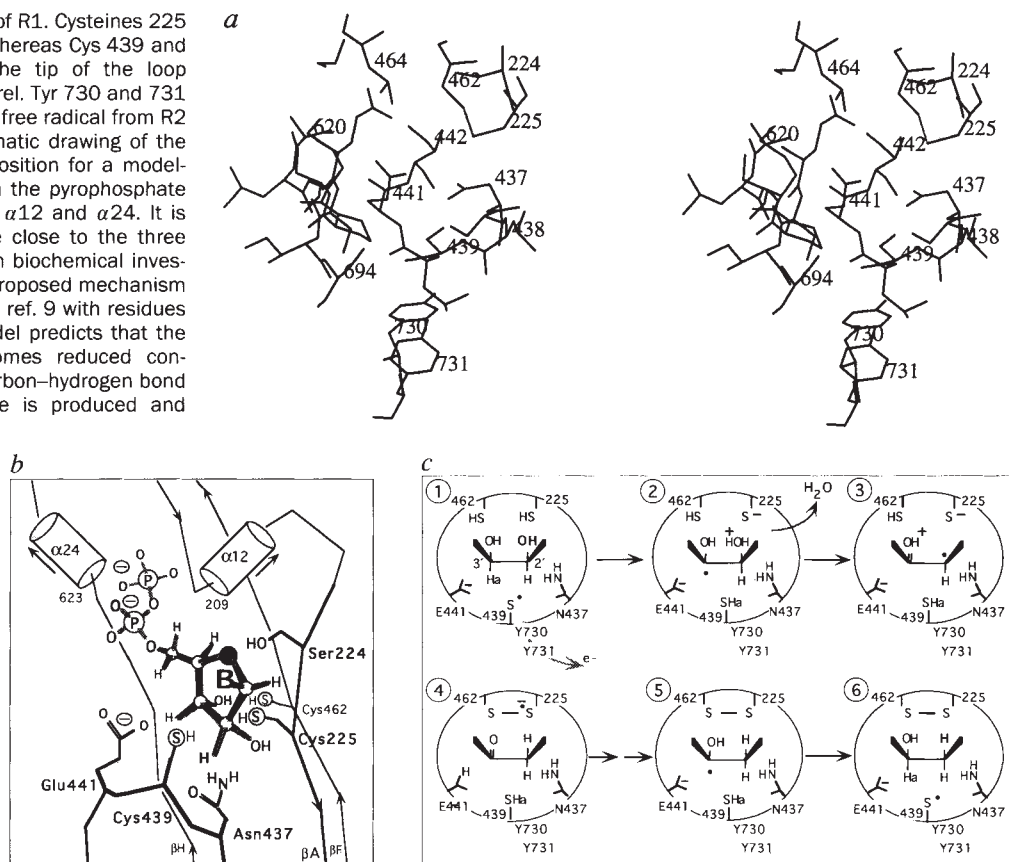


FIG. 4 a, Residues at the active site of R1. Cysteines 225 and 462 form a disulphide bridge, whereas Cys 439 and sequentially close residues form the tip of the loop inserted in the centre of the α/β barrel. Tyr 730 and 731 may participate in the transfer of the free radical from R2 to the active site Cys 439. b, Schematic drawing of the active-site region with a plausible position for a model-built substrate molecule placed with the pyrophosphate group at the amino ends of helices $\alpha 12$ and $\alpha 24$. It is then possible to position the ribose close to the three active-site cysteines, compatible with biochemical investigations. B, Substrate base. c, The proposed mechanism for substrate reduction adopted from ref. 9 with residues of the active site included. This model predicts that the one-electron-oxidized Cys 439 becomes reduced concomitantly with cleavage of the 3' carbon-hydrogen bond of the substrate. A deoxynucleotide is produced and Cys 439 is reoxidized to a thyl radical after a complex series of transformations involving loss of H_2O followed by numerous electron and proton transfers. Cysteines 225 and 462 at the top are reducing the substrate and become oxidized to a disulphide bridge. The radical is suggested to be transferred to the substrate via Tyr 731, Tyr 730 and Cys 439. Glu 441 is proposed to act as a base in the reaction.



nal barrel loop are stabilized by hydrogen bonds to main-chain nitrogen atoms of residues 441 and 442.

There are no conserved positively charged residues in the active site region. The closest positively charged residue at the active site is Arg 411 about 10 Å from the active site cysteines whereas the closest lysine residues (154 and 320) are further away, at the surface of the subunit. The most plausible phosphate-binding site, not far from the active-site cysteine residues, appears to be at the amino end of helices $\alpha 12$ and $\alpha 24$ (Fig. 4b). This arrangement is similar to other barrel structures that use short helices for phosphate binding²⁷. There appear to be two sulphate ions bound at the end of $\alpha 24$. Soaking of the crystals with the substrate cytidine diphosphate gives no significant differences in the electron density. The high sulphate concentration in the crystallization solution probably prevents CDP binding.

Structural similarities between reductases

The R1 sequences of the class I RNRs from different species (bacterial, mammalian and herpes virus origins) show weak but significant sequence similarities distributed along the whole polypeptide chain except for the first 70 residues where the similarities are weak. This strongly suggests that the R1 subunits from other species have the same overall three-dimensional structure as *E. coli* R1.

There is no obvious general similarity between the sequence of the class II enzyme ribonucleotide reductase from *Lactobacillus leichmannii* and the class I R1 proteins^{28,29}, but they do have striking similarities in a few regions. They have a similar pattern of active cysteines and the sequence of the internal loop-finger around Cys 439 of *E. coli* R1 has 5 out of 10 residues identical to a corresponding sequence in *L. leichmannii* (residues corresponding to Asn 437, Cys 439 and Glu 441 are present in this part). Furthermore, there is a 50% identity for the C-terminal 20 residues of the *E. coli* and *L. leichmannii* enzymes. This is a larger similarity than within the class I enzymes for this region. A third stretch of high similarity occurs around residue 600 in

E. coli located on the backside of the subunit far from the active site. Thus, although it appears that the overall structure of the *L. leichmannii* enzyme may be similar to the *E. coli* R1 protein, the enzymes have diverged widely from each other to the extent of even using different radical generating systems.

Binding of the R2 peptide

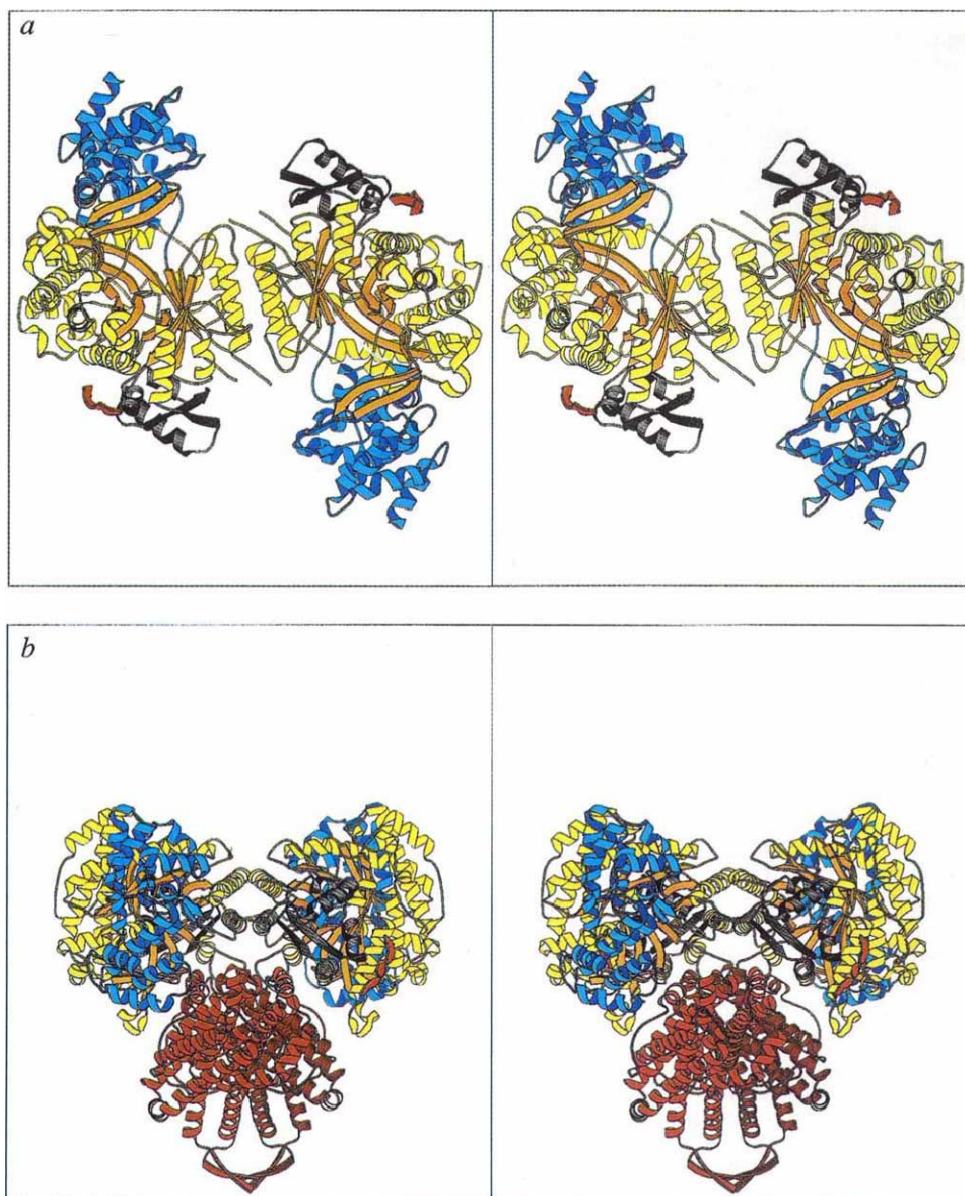
The C terminus of R2 is important for binding to R1 and peptides corresponding to the C terminus of R2 inhibit the formation of the holoenzyme complex. These interactions are species specific. The binding of carboxyl end peptides of R2 to R1 is of interest because of its possibilities as a basis for the design of species-specific antiproliferative drugs.

Electron density corresponding to the carboxyl end of the peptide of R2, which was added in the crystallization is visible in density maps in all three subunits in the asymmetric unit. The last residues of the peptide are best defined. The C terminus of the peptide is inserted in a shallow groove between helices $\alpha 13$ and $\alpha 1$ with contacts to the second half of the barrel and to the small domain (Fig. 3c). The residues surrounding the peptide-binding site are not conserved between species, reflecting the large deviations between the carboxyl peptides of R2 and the species-specific interaction between R1 and R2. The only similarity at the R2 carboxyl end of different species is that the last residue is a Leu or Phe, compatible with the observed hydrophobic binding site.

R1 dimer and a proposed interaction with R2

The R1 dimer is S-shaped (Fig. 5a) with a length of 110 Å and a width of 75 Å. The dimer interactions are formed by two helices of the α/β barrel, αA and αB and the symmetry-related pair. The shape of the R1 dimer is unusual and the dimer itself seems unstable. It is obvious that both subunits and domains can move with respect to each other which may form a structural basis for the allosteric regulation and catalysis. The shape of the R1 dimer is complementary to the upper part of the heart-shaped structure of the R2 dimer and a model of the complex can be

FIG. 5 a, The R1 dimer, where the two subunits are coloured as in Fig. 2a. The R2 peptides are coloured red. Three molecular dimers are arranged as hexamers in the crystals with three subunits in the asymmetric unit¹⁷. The subunits in the R1 dimer are related by crystallographic 2-fold axes in one of these hexamers but with non-crystallographic 2-fold axes in the others. Shown in the picture are two non-crystallographically related subunits. The main interactions between the subunits are formed by helices A and B in each α/β barrel. The interactions between the subunits consist mainly of van der Waals contacts between these two helices and only a few hydrogen bonds and charged interactions. An accessible surface area of 2,600 Å² is buried on dimer formation. The dimer interactions are thus weak and narrow in relation to the large size of the subunit. Loops 1, 2 and 3 from each subunit come close to each other, close to the 2-fold axis. They show weak density, indicating that they are flexible. These loops may obtain a fixed position in the R1–R2 complex to stabilize the holoenzyme. b, The R2 dimer (at the bottom) can be docked with the R1 dimer (on top) showing complementarity. The 2-fold axes of each dimer superimpose automatically. The proposed $\alpha_2\beta_2$ holoenzyme complex forms a compact entity with the R1 dimer on top of the R2 dimer. The best fit which is obtained when R2 is bound with its upper part of the heart to R1, is consistent with the location of conserved residues in this part of the R2. The C terminus of R2 can reach the binding site of the 20-mer peptide. The R1 molecule is coloured as in a. The R2 dimer and the peptides are in red.



formed with the 2-fold axes of the dimers superimposed (Fig. 5b). In this compact complex, R2 fits neatly into the cavities of R1 and the R1 dimer sits like a saddle on the top of the heart-shaped R2 dimer. The resulting complex appears to be considerably more stable than the R1 dimer, and there is still space for the substrate to pass into the active site.

The R1 structure and the mechanism of action

A detailed proposal for the mechanism of action of ribonucleotide reductase has been developed during the last decade and the last thorough description is depicted in ref. 9. The transformation of the substrate is initiated by the tyrosyl radical of R2. Originally, the active site was thought to be located at the R1–R2 interface with direct involvement of Tyr 122. However, the three-dimensional structure of R2 showed that Tyr 122 was buried about 10 Å inside the R2 protein which led to the suggestion that the radical is transferred to the active site of R1^{6,7}.

The R1–R2 model indicates how the transfer of the radical from R2 to R1 could occur. Part of such a pathway from Tyr 122 through Trp 48 to the active site on R1 was earlier suggested^{6,7}. The area around Trp 48 from both R2 subunits interacts in the complex model with the α/β barrels of R1 with the distance to the active site being about 25 Å. Any other modelling of the

complex will not bring Trp 48 or Tyr 122 closer to the active site. The structure of R1 shows that the active-site cysteines are in the inner parts of the substrate pocket. Because the barrel is very rigid it is not expected that conformational changes would bring the active-site cysteines closer to R2. Thus, one of the conclusions to be drawn from the structure of R1 is that the communication between Tyr 122 and the active site on R1 has to be a long-range radical transfer.

Cys 439 has been suggested to be the residue at the active site directly responsible for the radical initiation on the ribose of the substrate⁹. Which part of the R1 molecule that may participate in the electron transfer from Cys 439 has not yet been investigated by biochemical techniques. From our model of the holoenzyme complex, the two conserved tyrosine residues, 730 and 731, are located such that they may be involved in the radical transfer reaction. The OH of Tyr 730 is hydrogen bonded both to the sulphur of Cys 439 and OH of Tyr 731 and the side chains of the two tyrosines are in van der Waals contact to each other. The tyrosines are consecutive residues in a β -strand but a kink in the strand brings them together on the same side of the strand with unusual main-chain torsion angles for Tyr 730 (Figs 1 and 4a).

The remaining gap between Tyr 731 of R1 and Trp 48 of R2 may be bridged by Tyr 356 of R2 which was suggested from mutagenesis experiments to be involved in the electron transfer between R1 and R2³⁰. Unfortunately, the 35 last residues at the carboxyl end (including Tyr 356) were not visible in the R2 structure because of high flexibility^{6,7,31}. Tyr 356 is also not visible in the R1-peptide complex. However, if the visible part of the peptide and the last visible residue of R2 (340) are connected, Tyr 356 falls in the right area, but there are uncertainties in such a modelling because many residues are missing.

To be able to understand the specific role of the cysteines at the active site, a substrate was model-built (Fig. 4b). The pyrophosphate group of the substrate was positioned at the amino ends of α 12 and α 24. This part of the ribonucleotide is then bound in an internal pocket which is consistent with the fact that only nucleoside diphosphates are substrates. The ribose moiety of a nucleoside triphosphate would not be able to bind close to the active site cysteines at the same time as the pyrophosphates are bound at the phosphate-binding site. The ribose moiety of the diphosphate substrate can be bound such that O3' is at hydrogen-bonding distance to Glu 441, O2' to Asn 437 and O4' to Ser 224. The base of the substrate is positioned close to the conserved Gly 253 at the start of β B.

The first step in the proposed mechanism is a hydrogen removal of the 3'-hydrogen atom by a thiol radical at Cys 439⁹ (Fig. 4c). The positioning of Cys 439 is consistent with its suggested role in the activity of the enzyme because the ribose can be positioned with its 3' hydrogen directed towards Cys 439. The removal of this hydrogen is followed by a protonation of the 2'-hydroxyl, which may be achieved by Cys 225, which is close to O2'. Such a protonation helps leaving of a water molecule⁹. The resulting positive charge appearing at O3' may be stabilized by Glu 441.

The redox reaction is done by Cys 225 and Cys 462. The location of these two residues at adjacent strands forming a disulphide bridge in the structure is strong support for these as redox-active cysteines. Data from a reduced crystal (dithiothreitol (DTT) soaked) show no movement of Cys 225, but a rotation of Cys 462 which moves the sulphur away from Cys 225 and Cys 439. The redox-active cysteines and Cys 439 are located on opposite sides of the ribose ring, and should be able to participate directly in the reaction with the substrate. Ribonucleotide reductase is the only enzyme described in detail where cysteines can reduce a carbohydrate substrate. Usually cysteine residues reduce other thiol compounds by simple thiol exchange reactions as in glutathione reductase and thioredoxin

reductase^{32,33}. The disulphide bridge is in those cases reduced by flavins.

The active site cleft between the N terminal and barrel domains is not wide enough to allow a glutaredoxin or thioredoxin to enter and directly reduce the disulphide bridge between Cys 225 and Cys 462 which is formed after substrate reduction. Two additional cysteines (754 and 759) are involved in the reactivation of the enzyme^{22,24}. The last 24 residues at the carboxyl end, which includes these two cysteines, are not visible in the electron density map. This agrees with the hypothesis that the carboxyl end of R1 acts as a flexible shuttle of reducing equivalents from the surface of the molecule to the active site^{9,22}. The distance from the last visible residue (737) to Cys 225 is about 30 Å. Cys 225 is more accessible than Cys 462, suggesting it to be directly involved in the thiol exchange reaction when the disulphide bridge between Cys 225 and Cys 462 is reduced by cysteins 754 and 759. The arrangement of the flexible carboxyl end allows the enzyme to have a buried active site suitable for the substrate reduction, at the same time as to have a specific reduction system. The disordered carboxyl end also implies that the specificity for thioredoxin and glutaredoxin should be possible to investigate on peptides similar to the carboxyl end of R1.

Binding of effector molecules have been tried. Difference Fourier maps at low resolution of crystals soaked in dATP and crystals grown in the presence of dTTP give positive peaks at the amino end of α A close to the dimer interaction. This area is close to Cys 292 in the other subunit which has been labelled by nucleotide effector analogues³⁴. High-resolution difference Fouriers have been blurred by lack of isomorphism and these complexes have to be investigated further.

Discussion

The structure of R1 has demonstrated that Cys 225 and Cys 462 are indeed the redox-active cysteines. The close proximity of these residues to Cys 439 demonstrates that these three cysteines can act directly on the ribose of the substrate. The structure has further suggested that Ser 224, Asn 437 and Glu 441 are involved in ribose binding and transformation and that Tyr 730 and Tyr 731 participate in electron transfer between R1 and R2. The narrow active-site cleft in R1 excludes direct reduction of active-site cysteines by glutaredoxin or thioredoxin. Together with the disordered carboxyl end of the peptide this provides strong support for the theory of a flexible tail that transfers reducing equivalents from the surface to the active site by a thiol exchange reaction. □

Received 9 February; accepted 4 July 1994.

- Elledge, S. J., Zhou, Z. & Allen, J. B. *Trends biochem. Sci.* **17**, 119–123 (1992).
- Tuggle, C. K. & Fuchs, J. A. *J. Bact.* **172**, 1711–1718 (1990).
- Björklund, S., Skog, S., Tribukait, B. & Thelander, L. *Biochemistry* **29**, 5452–5458 (1990).
- Reichard, P. *Science* **260**, 1773–1777 (1993).
- Reichard, P. & Ehrenberg, A. *Science* **221**, 514–519 (1983).
- Nordlund, P., Sjöberg, B.-M. & Eklund, H. *Nature* **345**, 593–598 (1990).
- Nordlund, P. & Eklund, H. *J. molec. Biol.* **232**, 123–164 (1993).
- Thelander, L., Larsson, B., Hobbs, J. & Eckstein, F. *J. biol. Chem.* **251**, 1398–1405 (1976).
- Mao, S. S. et al. *Biochemistry* **31**, 9733–9743 (1992).
- Mao, S. S. et al. *Biochemistry* **31**, 9744–9751 (1992).
- Mao, S. S., Yu, G. X., Chalfoun, D. & Stubbe, J. *Biochemistry* **31**, 9752–9759 (1992).
- Holmgren, A. *J. biol. Chem.* **264**, 13463–13460 (1989).
- Moss, N. et al. *J. med. Chem.* **36**, 3005–3009 (1993).
- Cohen, E. A., Gaudreau, P., Brazeau, P. & Langlier, Y. *Nature* **321**, 441–443 (1986).
- Dutia, B. M., Frame, M. C., Subak-Sharpe, J. H., Clark, W. N. & Marsden, H. S. *Nature* **321**, 439–441 (1986).
- Clement, I., Sjöberg, B.-M. & Huang, C. Y. *Biochemistry* **30**, 5164–5171 (1991).
- Uhlén, U., Uhlén, T. & Eklund, H. *FEBS Lett.* **336**, 148–152 (1993).
- Zhang, K. Y. J. & Main, P. *Acta crystallogr.* **A46**, 377–381 (1990).
- Jones, T. A. in *Molecular Replacement* (ed. Dodson, E. J.) 91–105 (SERC, Daresbury, UK, 1992).
- Farber, G. K. & Petsko, G. A. *Trends biochem. Sci.* **15**, 228–234 (1990).
- Brändén, C.-I. *Curr. Opin. struct. Biol.* **1**, 978–983 (1991).
- Åberg, A. et al. *J. biol. Chem.* **264**, 12249–12252 (1989).
- Åberg, A. thesis, Univ. Stockholm (1993).
- Lin, A.-N.I., Ashley, G. W. & Stubbe, J. *Biochemistry* **26**, 6905–6909 (1987).
- Mao, S. S., Jokuston, M. I., Bollinger, J.-M. & Stubbe, J. *Proc. natn. Acad. Sci. U.S.A.* **86**, 1485–1489 (1989).
- Karlsson, M., Sahlin, M. & Sjöberg, B.-M. *J. biol. Chem.* **267**, 12622–12626 (1992).
- Wilmanns, M., Hyde, C. C., Davies, D. R., Kirschner, K. & Jansonius, J. N. *Biochemistry* **30**, 9161–9169 (1991).

- Booker, S. & Stubbe, J. *Proc. natn. Acad. Sci. U.S.A.* **90**, 8352–8356 (1993).
- Booker, S., Broderick, J. & Stubbe, J. *Biochem. Soc. Trans.* **21**, 727–730 (1993).
- Clement, I., Sjöberg, B.-M. & Huang, C. Y. *Biochemistry* **31**, 4801–4807 (1992).
- Lycksell, P.-O., Ingemarson, R., Davis, R., Gräslund, A. & Thelander, L. *Biochemistry* **33**, 2838–2842 (1994).
- Schulz, G. E., Schirmer, R. H., Sachsenheimer, W. & Pai, E. F. *Nature* **273**, 120–124 (1978).
- Prongay, A. J. & Williams, C. H. Jr. *J. biol. Chem.* **265**, 18968–18975 (1990).
- Eriksson, S., Sjöberg, B.-M., Jörnvall, H. & Carlquist, M. *J. biol. Chem.* **261**, 1878–1882 (1986).
- Blum, M., Metcalf, P., Harrison, S. C. & Wiley, D. C. *J. appl. Crystallogr.* **20**, 235–242 (1987).
- Nyberg, J. & Wonacott, A. J. in *The Rotation Method in Crystallography* 139–152 (North Holland, Amsterdam, 1977).
- Otwinowski, Z. *Denzo Manual* (Yale University, New Haven, 1991).
- Knight, S. thesis, Swedish Univ. Agricultural Sciences (1989).
- Terwillinger, T. C. & Eisenberg, D. *Acta crystallogr.* **A39**, 813–817 (1983).
- Otwinowski, Z. *MLPHARE, CCP4 Proc.* 80–88 (Daresbury Laboratories, Warrington, UK, 1991).
- Wang, B. C. *Meth. Enzym.* **115**, 90–112 (1985).
- Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. *Acta crystallogr.* **47**, 110–119 (1991).
- Jones, T. A. & Thirup, S. *EMBO J.* **5**, 819–822 (1986).
- Bricogne, G. *Acta crystallogr.* **A30**, 395–405 (1974).
- Brünger, T. A., Kuriyan, J. & Karplus, M. *Science* **235**, 458–460 (1987).
- Kraulis, P. J. *J. appl. Crystallogr.* **24**, 946–950 (1991).
- Presnell, S. R. & Cohen, F. E. *Proc. natn. Acad. Sci. U.S.A.* **86**, 6592–6596 (1989).

ACKNOWLEDGEMENTS. We thank B.-M. Sjöberg and her group for the supply of protein in earlier stages of the investigation and for the gift of recombinant bacteria, A. Jones for great help with the program O, P. Nordlund for early computational work, B.-M. Sjöberg, P. Reichard and L. Thelander for comments about the manuscript and T. Bergfors for linguistic corrections of the manuscript. This work was supported by grants from the Swedish Natural Science Research Council.

Selective activation of carbon-carbon bonds next to a carbonyl group

Masahiro Murakami, Hideki Amii
& Yoshihiko Ito

Department of Synthetic Chemistry, Faculty of Engineering,
Kyoto University, Yoshida, Kyoto 606-01, Japan

ORGANOMETALLIC complexes are used to effect a wide range of catalytic transformations in organic synthesis, such as the activation of C-H bonds^{1,2}. Carbon-carbon bonds, however, are generally unreactive towards transition metals under homogeneous conditions. C-C bond activation by a process of oxidative addition to soluble transition-metal complexes has been limited mostly to stoichiometric (not catalytic) reactions^{1,3-7,18} to highly strained substrates such as cyclopropane and cubane^{1,8-11} or to chelating ketones¹⁹. Here we present a synthetically useful process of selective C-C bond activation in which the C-C bond adjacent to a carbonyl group is opened by insertion of a soluble rhodium(I) complex. The resulting organometallic intermediate can be transformed to a variety of products in a way that regenerates the rhodium complex. We anticipate that this catalytic scheme will have considerable utility in organic synthesis.

We have found that Rh(I) can be inserted into a variety of C-C bonds next to a carbonyl group. In some cases this can lead to decarbonylation—removal of the carbonyl group—whereas in others a catalytic transformation can be effected. As an example of the former, cyclobutanone (**1** in Fig. 1) was treated with an equimolar amount of $((C_6H_5)_3P)_3RhCl$ under reflux for 41 h in toluene. Decarbonylation took place, giving quantitatively the corresponding cyclopropane **4**. The formation of **4** suggests that Rh(I) undergoes an insertion into the bond between carbonyl carbon and the α -carbon, giving the five-membered acylrhodium **2** in the initial step. Extrusion of the carbonyl group from the five-membered ring then occurs with migration onto rhodium to furnish the contracted rhodacycle **3**. Subsequent reductive elimination gives rise to the cyclopropane **4** together with the rhodium carbonyl complex.

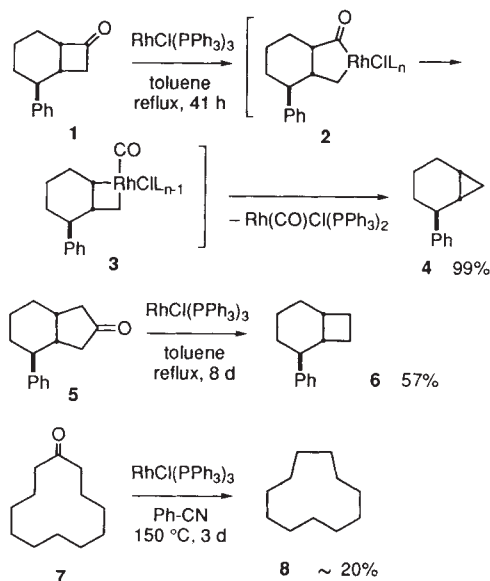


FIG. 1 The Rh-mediated decarbonylation of ketones.

like cyclopentanone **5** and cyclododecanone **7** (Fig. 1), also underwent the rhodium-mediated decarbonylation, although less efficiently, disclosing the general ability of Rh(I) to insert into the C-C bond α to a carbonyl group. Because a stoichiometric amount of $((C_6H_5)_3P)_3RhCl$ was required for the decarbonylation reaction, it is likely that the resulting rhodium carbonyl complex¹² fails to activate the C-C bond again. The formation of the stable rhodium carbonyl complex analogous to Vaska's complex may provide the driving force for the present decarbonylation reaction.

Next, the Rh-mediated selective activation of the C-C bond α to a carbonyl group was combined with a process of hydrogenolysis, thus leading to a catalytic reaction. When cyclobutanone **1** was treated under a hydrogen atmosphere (50 atm) with Rh(I) catalyst (10 mol%), prepared *in situ* from $[(cod)RhCl]_2$ and dppe, (cod is 1,5-cyclooctadiene, dppe is 1,2-bis(diphenylphosphino)ethane), alcohol **10** (Fig. 2) was produced in 87% isolated yield. It is likely that the aldehyde **9** is formed first by the addition of dihydrogen to **2'**. The following addition of dihydrogen to **9** gives the alcohol **10**. Actually, the use of 5 mol% of the rhodium catalyst afforded a mixture of **9** and **10** (~1:1) in 81% total yield. The formation of cyclobutanol **11** was not observed, indicating that oxidative addition of the C-C bond proceeds faster than direct hydrogenation of the carbonyl group under

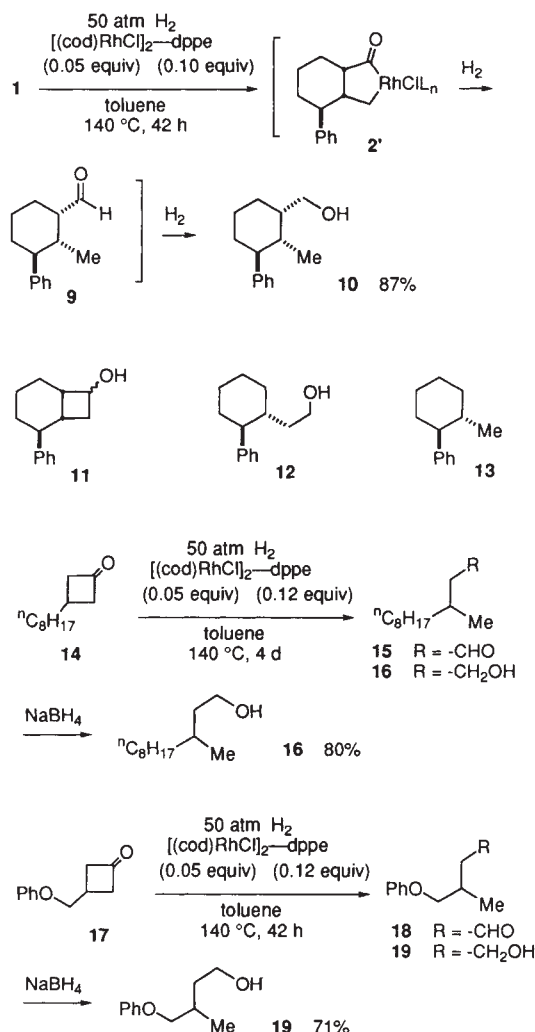


FIG. 2 The Rh-catalysed hydrogenolysis of the C-C bond α to a carbonyl group. (Abbreviations: cod, 1,5-cyclooctadiene; dppe, 1,2-bis(diphenylphosphino)ethane.)

these conditions. Moreover, alcohol **12** was not produced. This result shows that the insertion of Rh(I) took place selectively into the less-substituted C–C bond. The choice of the phosphine ligand is important for the product selectivity; when $((C_6H_5)_3P)_3RhCl$ was used, decarbonylation products (cyclopropane **4** (35%) and 1-methyl-2-phenylcyclohexane **13** (15%)) were obtained together with alcohol **10** (47%). Presumably 1-methyl-2-phenylcyclohexane **13** is formed by hydrogenation of the rhodacycle **3**. Other bidentate ligands like 1,3-bis-(diphenylphosphino)methane and 1,4-bis(diphenylphosphino)-butane also afforded a mixture of **10**, **4** and/or **13**.

Cyclobutanones **14** and **17** (Fig. 2) also underwent Rh(I)-catalysed cleavage of the α C–C single bond to afford mixtures of aldehydes and alcohols. After reduction with $NaBH_4$, 3-methylundecanol **16** (a building block in the synthesis of an insect pheromone)¹³ and 2-methyl-1,4-butanediol derivative **19** (a bifunctional isoprenoid building block)¹⁴ were isolated, respectively, in good yields (Fig. 2). Thus catalytic hydrogenolysis of C–C single bonds is accomplished via Rh(I)-mediated activation of C–C bond under a hydrogen atmosphere. The addition of dihydrogen together with relief of the structural strain of the cyclobutanone skeleton may make a large contribution to the driving force of this catalytic reaction.

A number of excellent methods are currently available for the synthesis of cyclobutanones^{15–17}. In particular, [2+2] cycloaddition reaction of alkenes with ketene derivatives provides a simple

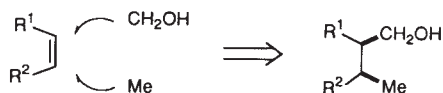


FIG. 3 Syn addition of methyl and hydroxymethyl groups to a C–C double bond.

and general approach to those compounds, often with control of the regio- and stereochemistry of ring substituents. Therefore, the present hydrogenolysis, coupled with preparation of cyclobutanones, achieves regio- and stereoselective *syn* 1,2-addition of methyl and hydroxymethyl groups to C–C double bonds (Fig. 3).

Although we have no experimental or theoretical result to provide a basis for argument about the mechanism of the insertion of Rh(I) into the α C–C bond, the reactions described here present the first example of practical synthetic transformations involving selective activation of C–C bonds by transition-metal complexes under homogeneous conditions, and demonstrate that the insertion of a transition metal into the C–C bond α to a carbonyl group may be a kinetically feasible fundamental process. Fine-tuning of the ligand set, the most favourable feature of homogeneous catalysts, and designing a reaction to gain a thermodynamic driving force would lead to the development of various kinds of transformations including those which can be applied to natural-product syntheses. Enantioselective hydrogenolysis of symmetrical cyclobutanones like **14** (Fig. 2) is one of the subjects worthy of further investigation. □

Received 9 June; accepted 4 July 1994.

- Crabtree, R. H. *Chem. Rev.* **85**, 245–269 (1985).
- Murai, S. *et al. Nature* **366**, 529–531 (1993).
- Gozin, M., Weisman, A., Ben-David, Y. & Milstein, D. *Nature* **364**, 699–701 (1993).
- Crabtree, R. H. & Dion, R. P. *J. chem. Soc., chem. Commun.* 1260–1261 (1984).
- Suggs, J. W. & Jun, C.-H. *J. Am. chem. Soc.* **106**, 3054–3056 (1984).
- Liebeskind, L. S., Baysdon, S. L., South, M. S. & Iyer, S. *Tetrahedron* **41**, 5839–5853 (1985).
- Hartwig, J. F., Andersen, R. A. & Bergman, R. G. *J. Am. chem. Soc.* **111**, 2717–2719 (1989).
- Bishop, K. C. *Chem. Rev.* **76**, 461–486 (1976).
- Tipper, C. F. H. *J. chem. Soc.* 2045–2046 (1955).
- Cassar, L., Eaton, P. E. & Halpern, J. J. *Am. chem. Soc.* **92**, 6366–6368 (1970).
- Periana, R. A. & Bergman, R. G. *J. Am. chem. Soc.* **108**, 7346–7355 (1986).
- Tsuji, J. & Ohno, K. *Tetrahedron Lett.* 3969–3971 (1965).
- Larcheveque, M., Sanner, C., Azerad, R. & Buisson, D. *Tetrahedron* **44**, 6407–6418 (1988).

- Schmid, R. & Hansen, H.-J. *Helv. chim. Acta* **73**, 1258–1275 (1990).
- Krepisk, L. R. & Hassner, A. *J. org. Chem.* **43**, 2879–2882 (1978).
- Bak, D. A. & Brady, W. T. *J. org. Chem.* **44**, 107–110 (1979).
- Gadwood, R. C., Mallick, I. M. & DeWinter, A. J. *J. org. Chem.* **52**, 774–782 (1987).
- Müller, E. & Segnitz, A. *Liebigs Ann. Chem.* 1583–1591 (1973).
- Suggs, J. W. & Jun, C.-H. *J. chem. Soc., chem. Commun.* 92–93 (1985).

SUPPLEMENTARY INFORMATION. Further details of experimental procedures for the reactions described, and of product characterization by 1H and ^{13}C NMR and elemental analysis, are available from Mary Sheenan at the London editorial office of *Nature*.

An efficient mimic of cytochrome P-450 from a zeolite-encaged iron complex in a polymer membrane

Rudy F. Parton*, Ivo F. J. Vankelecom*, Mark J. A. Casselman*, Cvetana P. Bezoukhanova†, Jan B. Uytterhoeven* & Peter A. Jacobs*

* K. U. Leuven, Department of Interface Chemistry, Centrum voor Oppervlaktechemie en Katalyse, Kardinaal Mercierlaan 92, B-3001 Heverlee (Leuven), Belgium

MANY attempts have been made to mimic the catalytic oxidative properties of the enzyme cytochrome P-450. For homogeneous systems¹ the mechanisms of oxidation can be readily determined but proper mimicry of the protein environment is difficult to achieve. Heterogeneous mimics have been designed that use organometallic complexes encapsulated in the supercages of zeolites^{2,3}, which enables control of selectivity and inhibition of auto-oxidation. But these systems do not show any mechanistic analogy with the enzymatic process, and the oxidation rates tend to be low. Here we report a composite catalytic system that achieves realistic mimicry of cytochrome P-450 as well as catalytic turnover rates that make the system industrially viable. Our catalyst incorporates iron phthalocyanine complexes encapsulated in crystals of zeolite Y, which are in turn embedded in a polydimethylsiloxane membrane. The polymer acts as a mimic of the phospholipid membrane in which cytochrome P-450 resides⁴, acting as an interface between two immiscible phases and avoiding the need for solvents or phase-transfer agents. This system oxidizes alkanes at room temperature at rates comparable to those of the enzyme⁵. The observation of a large kinetic isotope effect and the preferential oxidation of tertiary C–H bonds suggest close mechanistic similarities to the enzymatic process.

Previously reported turnovers for the room-temperature oxidation of alkanes with iodosobenzene as oxygen atom donor and zeolite-encapsulated iron phthalocyanine (FePc)^{6,7}, iron-tetramethylporphyrine⁸ or Mn-salen⁹ as catalysts, were below 10. With *t*-butylhydroperoxide (*t*-BHP), *n*-octane turnovers of 6,000 could only be obtained at low rates (0.2 turnovers min^{-1})¹⁰. Membrane-embedded metallo-porphyrins¹¹ showed up to 20 turnovers.

Our present catalyst, comprising iron phthalocyanine in zeolite Y (FePcY), was synthesized according to an optimized version of the metallocene method¹⁰ and was encapsulated in a polydimethylsiloxane (PDMS) membrane chosen for its high permeability and good affinity for the substrates used in the present catalytic reactions¹². The catalyst shows a 300-fold enhanced activity in alkane oxygenations compared to traditionally^{6,7} prepared zeolite-enclosed complexes. The composition of the FePcY–PDMS membrane is schematically represented

† On leave from University of Sofia, Faculty of Chemistry, Department of Organic Chemistry, J. Bouchier Avenue 1, BG-1126 Sofia, Bulgaria.

ted in Fig. 1. A cross-section through the membrane (scanning electron micrograph, Fig. 2) shows good dispersion of FePcY crystals in the membrane. No Pc needles could be detected.

We have observed room-temperature oxidation of cyclohexane with *t*-BHP at an average rate of 3.3 turnovers min^{-1} (Fig. 3, reaction E). Cyclohexane and *t*-BHP were chosen for their industrial relevance—they are the starting materials for production of adipic acid using a cheap and clean oxidant. The effect of the polymer membrane was studied in a batch liquid-phase reactor (Fig. 3, reactions A and B). The zeolite-Y crystals sequester the polar *t*-BHP and solvent at the expense of cyclohexane, resulting in lower rates of oxygenation and enhanced peroxide decomposition. A four-fold increase in activity is observed for the FePcY-PDMS system, which we attribute to the high hydrophobicity of the polymer, leading to an enhanced concentration of cyclohexane and a decreased concentration of peroxide in the PDMS matrix. Thus the polymer effectively acts as a 'solvent', adjusting the relative concentrations of both

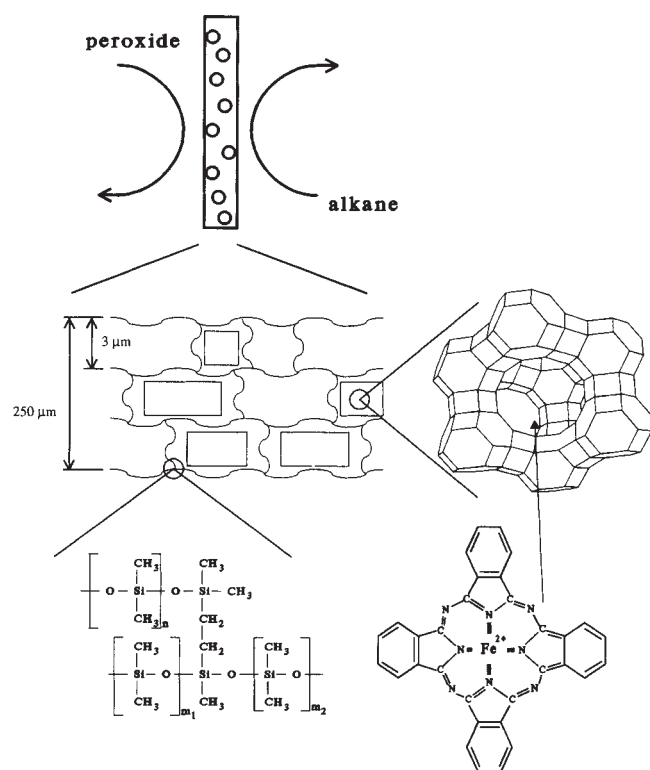


FIG. 1 Structural architecture of the FePcY-PDMS catalyst and its use in a counter-current reactor. FePcY is prepared according to an optimized version of an earlier procedure¹⁰. To NaY zeolite (Si/Al = 2.46) impregnated with ferrocene (1 molecule per supercage), an excess of 1,2-dicyanobenzene (8 molecules per supercage) was added. After heating at 453 K for 16 h under N_2 atmosphere, a blue-green crystalline solid is obtained which is purified by Soxhlet extraction with acetone, dimethyl formamide and acetone again. The FePcY has the faujasite X-ray diffraction pattern, devoid of phthalocyanine crystal reflections. Visible-near-infrared and infrared spectra indicate the presence of 0.5 FePc and 1.5 H_2Pc complexes per unit cell. The amounts of Fe and FePc, determined chemically and spectroscopically, respectively, point to the absence of significant quantities of uncomplexed iron. Identical (Si + Al)/Fe atomic ratios obtained chemically and with X-ray photoelectron spectroscopy (396 ± 40) point to quantitative encapsulation of FePc in the zeolite crystal. To dry FePcY (1.6 g) dispersed ultrasonically in methyl isobutylketone (4 g), 0.33 g of RTV 615B crosslinker (General Electric, New York) is added. After stirring this suspension for 2 h, 3.3 g of PDMS prepolymer (General Electric; RTV 615A) is added. The polymerizing mixture is cast on a glass plate as a 250- μm film and cured *in vacuo* at 423 K, giving FePcY-PDMS.

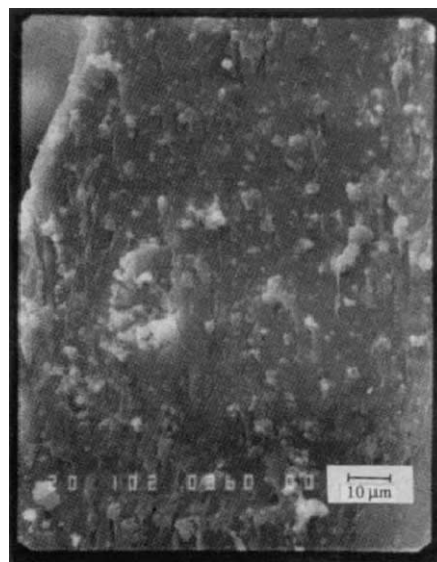


FIG. 2 Scanning electron micrograph of a cross-section of FePcY-PDMS membrane (thickness 65 μm ; 18.5 wt% FePcY).

reactants. Commonly used catalytic membranes¹³ consist either of thin, dense, noble-metal membranes or of catalytic components physically supported on porous membranes or chemically linked to the membrane polymer.

The oxygenation rate of FePcY can be improved in a fed-batch apparatus, where the addition of peroxide to the reaction mixture is controlled (Fig. 3, reactions C and D). The oxidation for FePcY is now three times faster than in batch conditions (Fig. 3, reaction A). As the membrane itself controls the sorption of the reactants, it is less influenced by the gradual addition of the peroxide. In the membrane reactor (Fig. 3, reaction E), a further increase in turnover (20-fold with respect to membrane-free catalysts, reaction A) and in efficiency occurs. The observed turnover rate of 3.3 min^{-1} is close to those reported for cytochrome P-450 (ref. 5). Moreover, the room temperature turnovers and peroxide efficiencies exceed by far those reported for Ti- β zeolites¹⁴ at 373 K with cyclododecane (Fig. 3, reaction F), which is an activated substrate compared to cyclohexane. Unfortunately, data are lacking that would allow comparison of FePcY-PDMS membranes with titanium-containing mesoporous molecular sieves^{15,16} such as Ti-MCM-41 and Ti-HMS. The perfect regeneration after reaction of our composite FePcY-membrane by vacuum drying, makes the system industrially viable.

Hydroxylation of aliphatic compounds⁵ by cytochrome P-450 and its model compounds involves oxidative cleavage of C-H bonds in the rate-determining step by a high-valent metallo-oxo species, followed by the rapid recombination of the metal-hydroxyl and alkyl radicals, avoiding formation of free radicals. The formation of Fe(v)-oxo species in this 'oxygen rebound mechanism'¹⁷, is proven spectroscopically for homogeneous mimics¹⁸. Evidence for hydrogen-atom abstraction by a symmetric linear transition state is provided by the existence of large kinetic isotope effect¹⁹⁻²¹. Based on the difference in zero-point-energy for C-H and C-D bond cleavage, room temperature kinetic isotope effects should not exceed 6.2 (ref. 22). Higher values are the result of secondary isotope effects, accounting for an increase of one or two units, and of the occurrence of hydrogen atom tunnelling during the attack on the C-H bond by the high-valent metal-oxo species²². The contribution of tunnelling is derived from the temperature dependence of the isotope effect.

Depending on the nature of oxygen-atom donor, metal and ligand structure, a broad range of kinetic isotope effects are reported for the oxidation of non-activated C–H bonds^{19–24}. A systematic summary is given in Fig. 4a for the cyclohexane/cyclohexane- d_{12} couple, including comparison with our new data. Free-radical oxidations¹⁹ have a kinetic isotope effect of below 2.5, whereas cytochrome P-450 and its 'real' homogeneous mimics (iron-porphyrins with NaOCl and iodosobenzene as oxygen-atom donors) have values of between 6.0 and 21.9 (refs 19–22, 24). This is understood in terms of a linear symmetric transition state^{19–21}. For kinetic isotope effects between 2.5 and 6, either a nonlinear or an asymmetric linear transition state is assumed²⁵.

For our FePcY membranes, a kinetic isotope effect of 9 ± 0.2 extrapolated to 0% conversion, was determined at room temperature. This also points to the existence of an electrophilic oxo-species, abstracting the hydrogen from cyclohexane by a symmetric linear transition state. As for FePcY, the difference in energy between hydrogen- and deuterium-atom abstraction ($E_{aD} - E_{aH} = 5.7 \text{ kJ mol}^{-1}$) is larger than the difference in zero-point energy (5.0 kJ mol^{-1}), some tunnelling will occur²². The ratio of the pre-exponential factors $A_H/A_D = 0.92$, which again confirms the existence of a symmetric linear transition state²⁵. It rejects the possibility of a free-radical mechanism and is totally in line with the 'oxygen rebound mechanism'.

It cannot be excluded, however, that small amounts of products are formed by a free-radical reaction, as traces of 1-chloro-adamantane are formed when adamantane is oxidized in dichloromethane. On the other hand, no *t*-butylperoxocyclohexane, *t*-butylcyclohexylether or bicyclohexyl is formed, indicating that oxidation by means of free alkyl radicals or homolytic scission of *t*-BHP is minor^{26,27} and does not lower the kinetic isotope effect substantially. The existence of adamantyl radicals shows that the reaction is not determined by other mechanisms as proton-, hydride- and proton-coupled electron-transfer mechanisms²⁸. The electrophilic character of the active iron-oxo intermediate is also supported by the inertness of chlorobutane

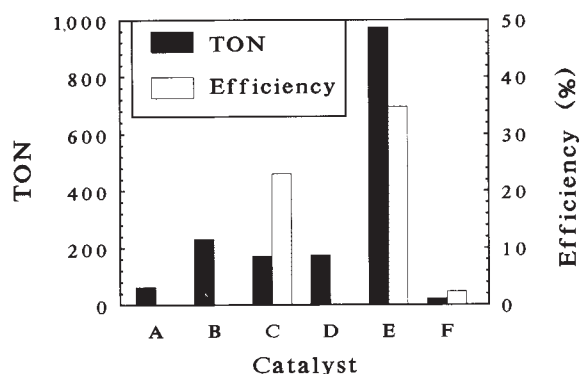


FIG. 3 Turnover numbers (TON) and peroxide yield for the oxidation of cyclohexane after 5 h reaction. Reactions A to E are carried out at room temperature and atmospheric pressure. Reaction A uses 0.096 g FePcY in a batch reactor with 300 ml of acetone as solvent, 40 mmol of a 7% aqueous *t*-BHP solution and 300 mmol of cyclohexane. Reaction B is identical to A with 0.32 g of FePcY-PDMS (0.096 g of FePcY) membrane pieces as catalyst. Reactions C and D are fed-batch reactions with 40 ml of acetone, 50 mmol of cyclohexane, and (C) 0.5 g of FePcY and (D) 1.67 g of FePcY-PDMS (0.5 g FePcY), aqueous *t*-BHP (70%) being added continuously at a rate of 4.38 mmol h^{-1} . Reaction E is done with 0.32 g of FePcY-PDMS (0.096 g of FePcY) in a counter-current membrane reactor with 40 mmol of *t*-BHP as a 7% solution in water and 300 mmol of cyclohexane. F refers to the batchwise cyclododecane oxidation at 373 K with $\text{Ti-}\beta^{14}$ as catalyst, using 16 mmol of cyclododecane, 62 mmol of aqueous H_2O_2 , 32 ml of acetone and 0.1 g of catalyst with 4.2 wt% of TiO_2 . Product separation is done by gas chromatography (GC) on a 50-m CP-Wax-52 CB capillary column (Chrompack, Middelburg).

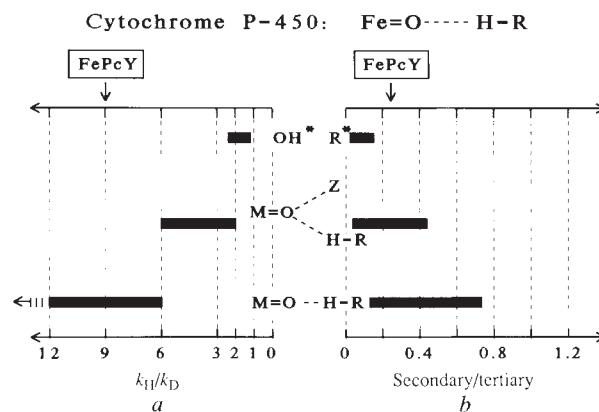


FIG. 4 Overview of room-temperature kinetic isotope effects in the oxidation of cyclohexane/cyclohexane- d_{12} mixtures (a) and relative reactivity of secondary and tertiary C–H bonds in adamantane (b). a, Competitive oxidation reaction of cyclohexane (25 mmol) with deuterated cyclohexane (25 mmol) on FePcY (0.5 g) at 298 K with *t*-BHP (50 mmol) and acetone (50 ml). The kinetic isotope effect, determined as the ratio of k_H and k_D , is extrapolated to 0% conversion. Product analysis is by GC and GC/mass spectrometry. Activation energy and pre-exponential factors are obtained by performing the experiments between 278 K and 323 K. b, Oxidation of adamantane (50 mmol) on 0.5 g FePcY at 298 K with 100 mmol *t*-BHP and 100 ml dichloromethane (see text for references).

for oxidation with FePcY, and the formation of 1-bromooctane-7-ol and -one from 1-bromooctane.

The reactivity order of primary, secondary and tertiary carbon atoms provides further evidence. Indeed, in the oxidation of adamantane (Fig. 4b) with FePcY, tertiary C–H bonds are found to react 11.4 times faster than secondary ones (secondary/tertiary ratio of 0.26; statistical value of 3). This is significantly higher than the ratios with Fenton²⁹ chemistry (0.05) and hydroxyl radicals²⁹ (0.14), but remains in the range required for oxo-chemistry⁵. By way of free hydroxyl radicals¹⁹, the ratios of reactivity for primary, secondary and tertiary carbon atoms is 1:5:12, respectively. This is in contrast with our FePcY system, as well as with cytochrome P-450 (ref. 30), where the oxidation of primary C–H bonds in *n*-alkanes is negligible.

Just as the lock-and-key mechanism of enzymes implies the selective diffusion of substrates to the active site, the zeolite framework exercises steric constraints on the active site of FePc, as does the protein mantle in cytochrome P-450. In the competitive oxidation of 2,3-dimethylbutane, 2-methylpentane and 3-methylpentane, the monobranched hydrocarbons are 10 times more reactive. Another aspect of the lock-and-key mechanism—steric constraints on the transition state—has been demonstrated by Tolman *et al.*⁶ for zeolite-enclosed complexes.

The newly developed concept of embedding catalysts in dense membranes is not restricted to the reported mimic of cytochrome P-450, but is applicable to any other reaction system involving immiscible reactants. Indeed, significant improvements on activity and efficiency can be expected for other promising oxidation catalysts, such as TS-1, Ti- β , Ti-MCM-41 and $[\text{Mn}(\text{Bpy})_2]^{2+}\text{-Na Y}$ (refs 31, 14, 15 and 32, respectively) when embedded in an appropriate dense membrane. It will make them even more competitive with existing industrial catalysts. $\square$

Received 16 May; accepted 20 July 1994.

- Meunier, B. *Chem. Rev.* **92**, 1411–1456 (1992).
- De Vos, D. E., Thibault-Starzyk, F., Knops-Gerrits, P. P., Parton, R. F. & Jacobs, P. A. *Macromol. Symp.* **80**, 157–184 (1994).
- Parton, R. F., De Vos, D. E. & Jacobs, P. A. in *Zeolite Microporous Solids: Synthesis, Structure and Reactivity* (eds Derouane, E. G. *et al.*) 555–578 (Kluwer Academic, Dordrecht, 1992).
- Ingelman-Sundberg, M. in *Cytochrome P-450: Structure, Mechanism, and Biochemistry* (ed. Ortiz de Montellano, P. R.) 119–160 (Plenum, New York, 1986).

5. Mansuy, D. & Battioni, P. in *Activation and Functionalization of Alkanes* Ch. 5 (ed. Hill, C. L.) 195–218 (Wiley, New York, 1989).
6. Tolman, C. A., Druliner, J. D., Nappa, M. J. & Herron, N. *Activation and Functionalization of Alkanes* Ch. 10 (ed. Hill, C. L.) 303–360 (Wiley, New York, 1989).
7. Ichikawa, M., Kimura, T. & Fukuoaka, A. *Stud. Surf. Sci. Catal.* **60**, 335–342 (1991).
8. Nakamura, M., Tatsumi, T. & Tominaga, H. *Bull. chem. Soc. Japan* **63**, 3334–3336 (1990).
9. Bowers, C. & Dutta, P. K. *J. Catal.* **122**, 271–279 (1990).
10. Parton, R. F., Uytterhoeven, L. & Jacobs, P. A. *Stud. Surf. Sci. Catal.* **59**, 395–404 (1991).
11. Groves, J. T. & Ungashe, S. B. *J. Am. chem. Soc.* **112**, 7796–7797 (1990).
12. Asada, T. *Membrane Science and Technology Series Vol. 1* 439–440 (Elsevier, Amsterdam, 1991).
13. Armor, J. N. *Appl. Catal.* **102**, 1–25 (1989).
14. Cambor, M. A., Corma, A. & Pérez-Pariente, J. *Zeolites* **13**, 82–87 (1993).
15. Tanev, P. T., Chibwe, M. & Pinnavaia, T. J. *Nature* **368**, 321–323 (1994).
16. Corma, A., Navarro, M. T. & Pérez-Pariente, J. *J. chem. Soc., chem. Commun.* 147–148 (1994).
17. Groves, J. T. *J. chem. Educ.* **62**, 928–931 (1985).
18. Penner-Hahn, J. E. et al. *J. biol. Chem.* **298**, 12761–12764 (1983).
19. Khenkin, A. M. & Shilov, A. E. *New J. Chem.* **13**, 659–667 (1989).
20. Sorokin, A. B. & Khenkin, A. M. *J. chem. Soc., chem. Commun.* 45–46 (1990).
21. Sorokin, A. B., Robert, A. & Meunier, B. *J. Am. chem. Soc.* **115**, 7293–7299 (1993).
22. Bell, R. P. *Chem. Soc. Rev.* **12**, 513–544 (1974).
23. Barton, D. H. R., Doller, D. & Geletti, Y. V. *Tetrahedron Lett.* **32**, 3811–3814 (1991).
24. Groves, J. T. & Nemo, T. E. *J. Am. chem. Soc.* **105**, 6243–6248 (1983).
25. Kwart, H. *Accs chem. Res.* **15**, 401–408 (1982).
26. Leishing, R. A., Kim, J., Pérez, M. A. & Que, L. Jr. *J. Am. chem. Soc.* **115**, 9524–9630 (1993).
27. Ménage, S. et al. *Inorg. Chem.* **32**, 4766–4773 (1993).
28. Che, C.-M., Ho, C. & Lau, T.-C. *J. chem. Soc., Dalton Trans.* 1259–1263 (1993).
29. Nam, W. & Valentine, J. S. *New J. Chem.* **13**, 677–682 (1989).
30. Ullrich, V. *Top. curr. Chem.* **83**, 67–104 (1979).
31. Huybrechts, D. R. C., De Bruycker, L. & Jacobs, P. A. *Nature* **345**, 240–242 (1990).
32. Knops-Gerrits, P.-P., De Vos, D., Thibault-Starzyk, F. & Jacobs, P. A. *Nature* **369**, 543–546 (1994).

ACKNOWLEDGEMENTS. We thank the Belgium government for support under the IUAP-PAI scheme. R.F.P. and J.F.J.V. thank the Belgium National Fund of Scientific Research (NFWO) for support, as post-doctoral researcher and research assistant respectively. C.P.B. is grateful to K. U. Leuven for a senior research Fellowship.

The role of stratospheric ozone in modulating the solar radiative forcing of climate

J. D. Halgh

Space and Atmospheric Physics Group, The Blackett Laboratory, Imperial College of Science, Technology and Medicine, London SW7 2BZ, UK

MANY recent studies have reported an apparent correlation between solar activity and the Earth's climate on the timescale of the 11-year solar cycle^{1–4} and over longer periods^{5–10}, but to date no physical mechanism has been proposed that can satisfactorily explain the observations. In general, it has been assumed that changes in total solar irradiance outside the atmosphere will be reflected in proportionately equal changes at the tropopause (from where the influence on climate is determined). Here I present results from a two-dimensional radiative-chemical-transport model which show that the spectral composition of the solar variations and the photochemical production of stratospheric ozone together lead to a highly nonlinear relationship between the extraterrestrial and cross-tropopause solar radiative flux. Because of this relationship, at middle to high latitudes in the winter hemisphere less solar radiation reaches the troposphere during periods of higher solar activity. The consequent change in latitudinal temperature gradient also affects infrared radiative forcing and potentially planetary-wave activity. The general mechanism proposed here may explain some features of the observed correlations between solar variability and climate.

Studies of the effects of solar variability on the Earth's atmosphere have been divided into two distinct areas. The first are those^{11–18} concerned with changes in the structure and chemical composition of the middle atmosphere. Measurements^{19–21} of the spectral composition of solar radiation reveal that, although the amplitude of the change in total solar irradiance between the minimum and maximum of the solar cycle is less than 0.1%,

this conceals much larger variations at the ultraviolet end of the spectrum. This ultraviolet radiation penetrates the mesosphere and upper stratosphere so that a response to the solar cycle has been demonstrated in the photochemistry at these altitudes with small (less than 2%) changes in total ozone column. The second area of research has been concerned with the contribution of changes in total solar irradiance to climate forcing. As the globally averaged net incoming solar radiation is $\sim 240 \text{ W m}^{-2}$, it has been assumed^{17,22,23} that a 0.1% increase in irradiance will result in a 0.24 W m^{-2} increase in climate forcing. In the work reported here, the two areas are brought together in an attempt to make a more precise assessment of changes in total solar flux at the tropopause by taking into account the effects of the intervening middle atmosphere.

A fully interactive two-dimensional radiative-chemical-transport model of the atmosphere was used; this model was developed from that of Harwood and Pyle²⁴. It is a classical eulerian model which extends from pole to pole with a resolution of 9.5° , and from the surface to a height of $\sim 95 \text{ km}$ with a resolution of $\sim 3.5 \text{ km}$. It contains detailed treatments of atmospheric chemistry and radiation²⁵ with the photolysis cross-sections and the top-of-atmosphere solar irradiance taken from WMO recommendations²⁶. Two model runs have been carried out for this study. The first is a control run representing a position of minimum activity in the 11-year solar cycle, whereas in the second the fluxes have been increased to represent solar maximum, using the measurements of Heath and Schlesinger¹⁹ between wavelengths λ of 168 and 300 nm, and the estimates of Lean²⁷ between 300 and 730 nm (Fig. 1). Although the values in the latter region are not known with any precision, their integral must be consistent with measurements of variation in total solar irradiance which indicate an increase of 0.08% between solar minimum and solar maximum²⁸. As they are not involved with ozone formation (although they are absorbed by ozone), any likely changes at these wavelengths will not significantly affect the conclusions of this paper (see later for discussion on $\lambda > 730 \text{ nm}$).

The difference in diurnally averaged solar flux between solar maximum and solar minimum as a function of height and wavelength is shown for circumstances appropriate to 26 December for latitudes 19° S and 57° N in Fig. 2a and b, respectively. Both show the increases impinging on the top of the atmosphere, though these are larger in absolute terms at the more equatorial latitude. At $\lambda < 242 \text{ nm}$ these fluxes force additional photodissociation of molecular oxygen, and the resulting atomic oxygen combines with O_2 to produce higher concentrations of ozone. This in turn absorbs more radiation especially between 200 and 350 nm but also, importantly, across the visible portion of the spectrum. The result of this additional absorption is that at 19° S the solar irradiance reaching the lower atmosphere is reduced in the region $205 \text{ nm} < \lambda < 320 \text{ nm}$ and intermittently in $560 \text{ nm} < \lambda < 670 \text{ nm}$, and increased elsewhere, while at 57° N , because of the longer atmospheric pathlength traversed and therefore ozone optical depth, it is reduced over a wider spectral region.

The effect on total (spectrally integrated) solar flux as a function of latitude and height is shown in Fig. 3a and on ozone concentration in Fig. 3b. The larger incident solar flux has, indeed, produced higher concentrations of ozone at all altitudes with results very similar to those of other recent two-dimensional model studies^{16,18} giving an increase in total column ozone of between 1.2 and 1.8%. As a result of the ozone increases, the full effect of the larger incident flux is not felt lower in the atmosphere; at the 100-mbar level in summer mid-latitudes the flux increase is only 0.10 W m^{-2} , or 0.025%, compared with an increase in incident radiation of 0.37 W m^{-2} , or 0.08% of the total solar input. Poleward of 45° N because of the extra absorption there is actually less solar flux at 100 mbar at solar maximum than at solar minimum; for example a 0.05% reduction at 65° N .

Consistent with other two-dimensional model studies^{16,18} stratospheric temperatures have also increased (Fig. 3c), and in response to these changes in temperature and ozone concentration the fluxes of thermal infrared radiation are also amended (Fig. 3d). The higher temperatures lead to greater emission mainly by CO₂, H₂O and O₃, whereas increases in ozone concentration will also tend to increase the radiance emitted in the 9.6- μ m band; these result in an increased downward flux of the order of 0.05 W m⁻² at 100 mbar.

In order to assess the implications of these findings for calculations of the solar radiative forcing of climate over the 11-year cycle, we must consider changes in radiance at the tropopause²⁹. It is clear that the assumption of a uniform 0.08% increase in solar radiation not only overestimates the direct solar effect everywhere, but is in the wrong sense at winter middle to high latitudes. The effect of changes in the downward longwave radiance, however, is to increase the flux at all latitudes with the result that at 100 mbar the net radiative forcing decreases from ~ 0.2 W m⁻² at the summer pole fairly uniformly to ~ 0.05 W m⁻² at the winter pole. It is not within the scope of the present work to predict the impact of this form of radiative forcing on climate but the latitudinal gradients predicted here, which are very different to those used previously especially near the winter pole, may result in the modulation of planetary-wave activity and stratospheric wind structure which have been suggested by observations^{4,30}.

It is also worth noting that the magnitude of the changes in radiative fluxes are of the same order as those discussed by Ramaswamy *et al.*³¹ in the context of the radiative forcing of climate by greenhouse gases over a decadal timescale, which will tend to make extraction of the different components from data records rather difficult.

In the work described earlier in this Letter, solar radiation at $\lambda > 730$ nm, which in energy terms comprises about half of the total, has not been included. If there were a uniform increase in solar irradiance between solar minimum and maximum across this part of the spectrum, then the effects of stratospheric ozone on the cross-tropopause flux would be proportionately reduced. Lean²⁷, however, estimates that although there is an increase from solar minimum to solar maximum of $\sim 0.04\%$ between 750 and 1,100 nm (a spectral region covering about a quarter of the total incoming irradiance) there is a compensating reduction of up to 0.03% in most of the remaining quarter. It appears, therefore, that changes in this half of the solar spectrum are unlikely to contribute significantly to changes in the radiative forcing of climate on the solar-cycle timescale. Further work needs to be done, however, to check this assumption, as there may be radiative-chemical feedbacks involving those gases with absorption bands in the near infrared.

Another factor which has been ignored in this work is the potential for changes in tropospheric composition and temperature, in response to the variation in solar fluxes, to affect the upward infrared flux at the tropopause, the therefore the net radiative forcing.

It is clear from the above that the solar radiative forcing of climate depends crucially on the spectral composition of the solar variability. The large observed changes at ultraviolet wavelengths cause

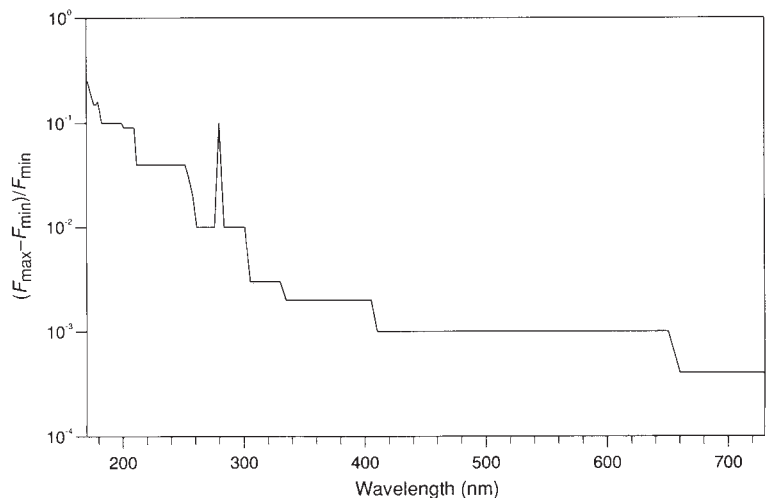


FIG. 1 Assumed fractional increase in incoming solar flux F from solar minimum ($F_{\min}$) to solar maximum ($F_{\max}$) as a function of wavelength.

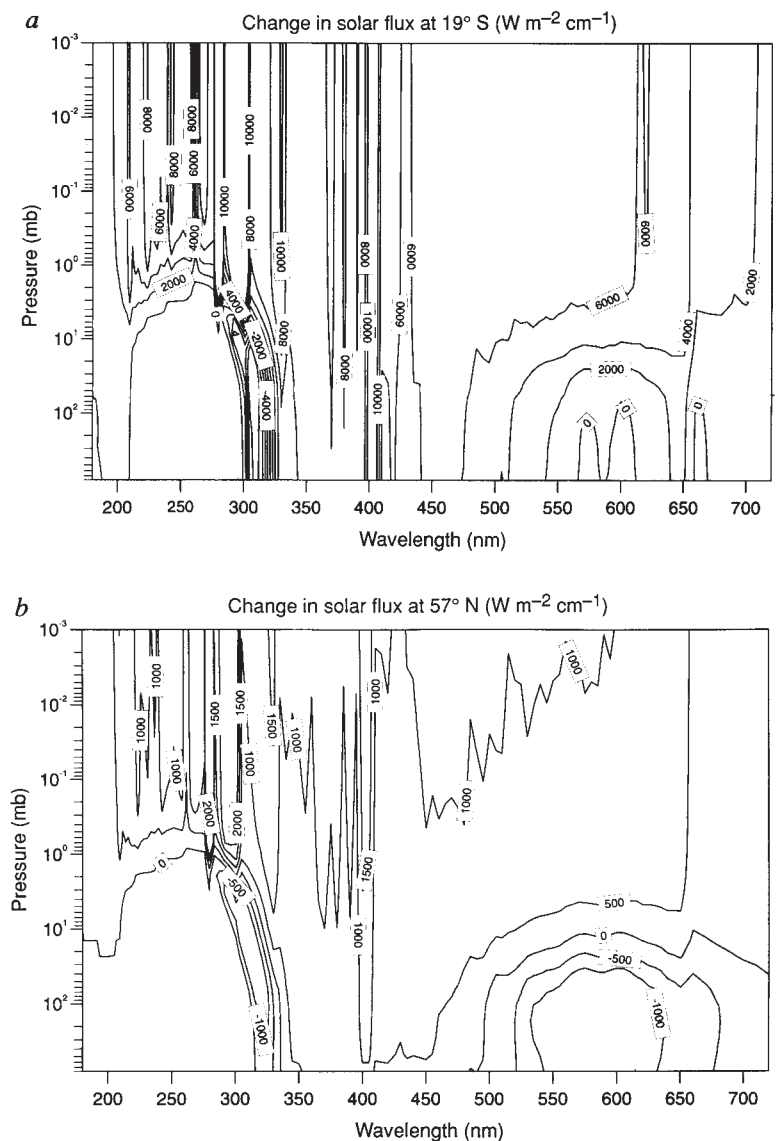


FIG. 2 Change in solar flux between solar maximum and solar minimum as a function of height (represented by atmospheric pressure) and wavelength on 26 December at 19° S (a) and 57° N (b).

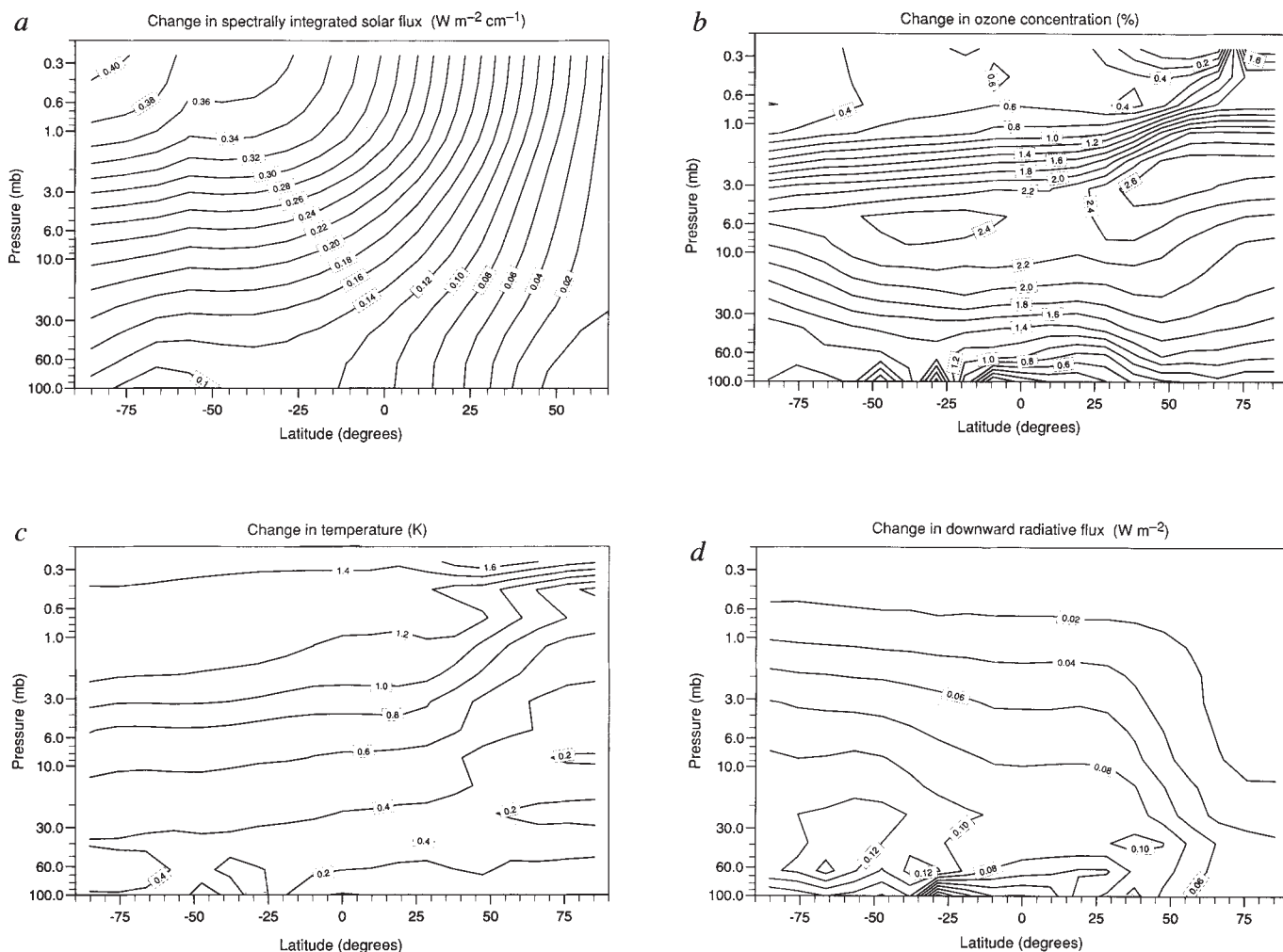


FIG. 3 a–d, Change in four parameters between solar maximum and solar minimum on 26 December as a function of height and latitude; a, spectrally integrated solar flux ($\lambda < 730$ nm); b, ozone concentration;

c, temperature, and d, spectrally integrated downward thermal infrared flux. (Note that positive latitude is N.)

variations in ozone concentration which can mask the lower atmosphere from the variations in the overall flux and even, at winter middle to high latitudes, reduce the radiative forcing at solar maximum relative to its solar minimum value. The degree of this masking depends sensitively on solar variability in the wavelength regions $340 \text{ nm} < \lambda < 500 \text{ nm}$ and $\lambda > 750 \text{ nm}$ which are little affected by ozone. The ozone and concomitant temperature changes in the lower stratosphere also affect the longwave radiation field.

The mechanism discussed here may explain some of the observed correlations between atmospheric parameters and solar variability. With regard to the correlation with solar-cycle length⁶, one direction of research might be to study how solar spectral variability is related to cycle length. This work does demonstrate that simple energy-balance ideas of solar forcing could well be misleading, and that there is a clear requirement for monitoring the entire solar spectrum over periods of decades. □

Received 12 January; accepted 1 July 1994.

1. Labitzke, K. *Geophys. Res. Lett.* **14**, 535–537 (1987).
2. Labitzke, K. & Van Loon, H. *J. Clim.* **2**, 554–565 (1989).
3. Van Loon, H. & Labitzke, H. *J. Clim.* **3**, 827–837 (1990).
4. Labitzke, K. & Van Loon, H. *Ann. Geophys.* **11**, 1084–1094 (1993).
5. Reid, G. C. *J. geophys. Res.* **96**, 2835–2844 (1991).
6. Friis-Christensen, E. & Lassen, K. *Science* **254**, 698–700 (1991).
7. Lean, J., Skumanich, A. & White, O. *Geophys. Res. Lett.* **19**, 1591–1594 (1992).
8. Kelly, P. M. & Wigley, T. M. L. *Nature* **347**, 460–462 (1990).
9. Kelly, P. M. & Wigley, T. M. L. *Nature* **360**, 328–330 (1992).
10. Schlesinger, M. E. & Ramankutty, N. *Nature* **360**, 330–333 (1992).
11. Penner, J. E. & Chang, J. S. *Geophys. Res. Lett.* **5**, 817–820 (1978).
12. Brasseur, G. & Simon, P. C. *J. geophys. Res.* **86**, 7343–7362 (1981).
13. Garcia, R., Solomon, S., Roble, G. R. & Rusch, D. W. *Planet. Space Sci.* **32**, 411–423 (1984).
14. Hood, L. L. *J. geophys. Res.* **91**, 5264–5276 (1986).
15. Chandra, S. *Geophys. Res. Lett.* **18**, 837–840 (1991).
16. Wuebbles, D. J., Kinnison, D. E., Grant, K. E. & Lean, J. *J. Geomagn. Geoelect.* **43** (suppl. p 22), 709–718 (1991).

17. Kodera, K., Chiba, M. & Shibata, K. *Geophys. Res. Lett.* **18**, 1209–1212 (1991).
18. Huang, T. Y. W. & Brasseur, G. P. *J. geophys. Res.* **98**, 20413–20427 (1993).
19. Heath, D. F. & Schlesinger, B. M. *J. geophys. Res.* **91**, 8672–8682 (1986).
20. Lean, J. *Science* **244**, 197–200 (1989).
21. Cebula, R. P., DeLand, M. T. & Schlesinger, B. M. *J. geophys. Res.* **97**, 11613–11620 (1992).
22. Hansen, J. E. & Lacis, A. A. *Nature* **346**, 713–719 (1990).
23. Wigley, T. M. L. & Raper, S. C. B. *Geophys. Res. Lett.* **17**, 2169–2172 (1990).
24. Harwood, R. S. & Pyle, J. A. *Q. Jl R. met. Soc.* **101**, 723–748 (1975).
25. Law, K. S. & Pyle, J. A. *J. geophys. Res.* **98**, 18377–18400 (1993).
26. *Atmospheric Ozone 1985* (Report No. 16, World Meteorological Organisation, Geneva, Switzerland, 1986).
27. Lean, J. *Rev. Geophys.* **29**, 505–535 (1991).
28. Willson, R. C. & Hudson, H. S. *Nature* **332**, 810–812 (1988).
29. Shine, K. P., Derwent, R. G., Wuebbles, D. J. & Morcrette, J.-J. in *Climate Change: The IPCC Scientific Assessment* (eds Houghton, J. T. et al.) 41–68 (Cambridge Univ. Press, 1990).
30. Kodera, K. *Geophys. Res. Lett.* **18**, 1023–1026 (1991).
31. Ramaswamy, V., Schwarzkopf, M. D. & Shine, K. P. *Nature* **355**, 810–812 (1992).

ACKNOWLEDGEMENTS. I am grateful to C. Winchester for help with the diagrams. This work was funded by the UK SERC and NERC.

Minimal effects of UVB radiation on Antarctic diatoms over the past 20 years

A. McMinn*, H. Heijnis† & D. Hodgson‡

* Antarctic CRC and Institute of Antarctic and Southern Ocean Studies, University of Tasmania, Box 252C, Hobart 7001, Tasmania, Australia

† Environmental Radiochemistry Laboratory, Australian Nuclear Science and Technology Organisation, Private Mailbag 1, Menai, New South Wales 2234, Australia

‡ Department of Plant Science, University of Tasmania, Box 252C, Hobart 7001, Tasmania, Australia

IT HAS been suggested¹⁻³ that increased springtime UVB radiation caused by stratospheric ozone depletion is likely to reduce primary production and induce changes in the species composition of Antarctic marine phytoplankton. Experiments conducted at Arthur Harbour in the Antarctic Peninsula revealed a reduction in primary productivity at both ambient and increased levels of UVB (ref. 4). Laboratory studies have shown that most species in culture are sensitive to high UVB levels, although the level at which either growth or photosynthesis is inhibited is variable^{5,6}. Stratospheric ozone depletion, with resultant increased springtime UVB irradiance, has been occurring with increasing severity since the late 1970s. Thus the phytoplankton community has already experienced about 20 years' exposure to increasing levels of UVB radiation. Here we present analyses of diatom assemblages from high-resolution stratigraphic sequences from anoxic basins in fjords of the Vestfold Hills, Antarctica. We find that compositional changes in the diatom component of the phytoplankton community over the past 20 years cannot be distinguished from long-term natural variability, although there is some indication of a decline in the production of some sea-ice diatoms. We anticipate that our results are applicable to other Antarctic coastal regions, where thick ice cover and the timing of the phytoplankton bloom protect the phytoplankton from the effects of increased UVB radiation.

Most interest in the effects of UVB radiation on Antarctic algae has centred on phytoplankton. This group of organisms, which is dominated by diatoms, occupies a pivotal position in the Southern Ocean food web. The phytoplankton at the ice-edge zone in spring are potentially susceptible to damage by ultraviolet radiation because here they are retained near the surface in a shallow lens of less saline water induced by the melting sea ice⁷. This occurs during spring when the area is also experiencing the maximum effects of enhanced UVB. Antarctic coastal phytoplankton are unlikely to be affected by increased UVB levels because the sea ice remains until December–January, by which time ambient UVB levels are relatively unaffected by ozone depletion.

In addition to the phytoplankton, sea-ice algae also make a significant contribution to the annual primary production of the Southern Ocean. Within the pack ice, infiltration, interior and bottom algal communities can be distinguished. In coastal areas much of the biomass is contributed by a mat and strand community, also largely comprised of diatoms, that grows on the bottom of the fast-ice⁸ each spring. Biomass of the overlying

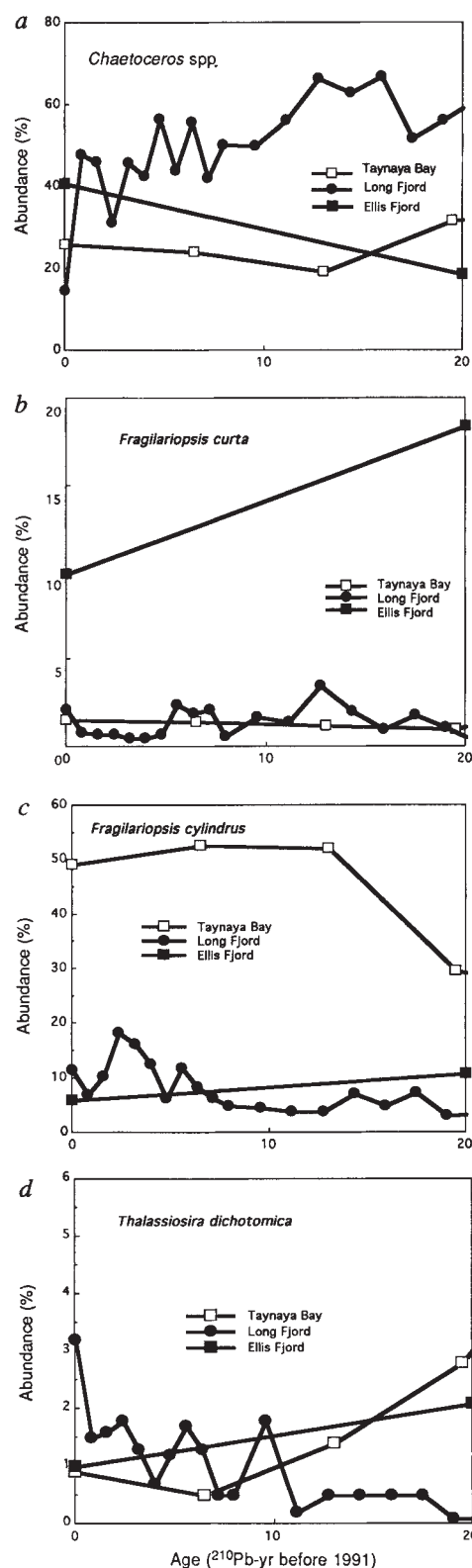


FIG. 1 a–d, Relative abundance of dominant phytoplankton taxa in sediment cores of the Vestfold Hills. a, *Chaetoceros* spp. vegetative cells (ratio of *Chaetoceros* vegetative cells to resting cysts remains approximately constant); b, *Fragilariopsis curta*; c, *Fragilariopsis cylindrus*; d, *Thalassiosira dichotomica*. Data points from Taynaya Bay and Ellis Fjord were taken at 1-cm intervals, data points from Long Fjord were taken at 0.5-cm intervals.

METHODS. The relative abundance of all common diatom species was calculated from a count of 1,200 frustules from each subsample. Only

those species with an abundance of >1% were considered further, as fluctuations in abundance below this level are subject to significant counting errors. The age of the samples was determined by estimating the sedimentation rate from ²¹⁰Pb analysis. The application of ²¹⁰Pb/²²⁶Ra as a tool for measuring sedimentation rates around the Antarctic continent is more difficult than elsewhere as the atmospheric fall of ²¹⁰Pb in this area is extremely low and the production rate of ²¹⁰Pb in the Southern Ocean is relatively high¹⁸. As a result excess ²¹⁰Pb contributes very little to the total ²¹⁰Pb in the core samples.

congelation ice is considerably lower than in these bottom communities (A.M., unpublished data). Many sea-ice algae are extremely shape-adapted⁹ and thus potentially susceptible to an increase in UVB levels. The fast-ice near the continent is relatively transparent, and transmits ~10% of UVB in October¹⁰ compared with only 0.25% in late November¹¹. Trondahl and Buckley¹¹ have estimated a 20-fold increase in the under-ice UVB irradiance in October due to ozone depletion in a period of relatively high transmittance. Snow cover on the ice has a major effect on light transmission to the water column, but in many coastal areas (for example, the Vestfold Hills in eastern Antarctica) the ice is often kept free of snow by strong katabatic winds.

The fjords of the Vestfold Hills are well positioned for an investigation of the response of a diatom community to increasing springtime UVB levels. They experience significant exchange of surface waters with the adjacent coastal zone but are sufficiently restricted to contain permanently stratified, anoxic basins that preserve high-resolution unbioturbated stratigraphic sequences. Three sediment cores, each ~40 cm long, were taken with a Glew corer from anoxic basins in Ellis Fjord, Long Fjord and Taynaya Bay in the Vestfold Hills in November 1992. The cores were taken from widely separated sites to minimize the influence of local climate and sea-ice effects. They were extruded and sectioned into either 1.0 or 0.5 cm intervals in the field. Subsamples were then taken for diatom analysis and ²¹⁰Pb sedimentation rate determination. Sedimentation rates, which were used to establish the time framework of deposition within the cores, were calculated using the constant initial concentration

method^{12,13}. The present interpretation of changes to algal communities in the Vestfold Hills is based on the investigation of diatoms alone, as the siliceous frustules of these organisms are readily fossilized and are consequently well preserved.

Interpretation of the effects of enhanced UVB levels on the diatom community over the past 20 years is based on changes in single-species abundance, species richness and assemblage diversity in the sediment cores. Most taxa show considerable fluctuations in abundance over relatively short time periods. An examination of the relative abundance of four of the most common phytoplankton taxa in Ellis Fjord¹⁴ (that is, *Chaetoceros* spp., *Fragilariopsis cylindrus*, *Fragilariopsis curta* and *Thalassiosira dichotomica*, which are abundant in most of the cores) shows that no consistent change in abundance has occurred across the Vestfold Hills over the past 20 years (Fig. 1). Investigation of changes in the abundance of these species over longer time periods shows fluctuations of a greater magnitude than are seen within the past 20 years (Fig. 2). This implies that the recent changes in abundance fall within the limits of natural variability, and that factors other than UVB are responsible for most of the changes in the phytoplankton community.

An examination of the three most abundant taxa from the sea-ice bottom mat and strand communities of Ellis Fjord¹⁴ (that is, *Entomoneis kjellmannii*, *Berkeleya rutilans* and *Nitzschia stellata*) in Taynaya Bay reveals a possible decline in abundance in the cores over the past 20 years. Low abundances preclude a more definitive interpretation (Fig. 3) and it is possible that

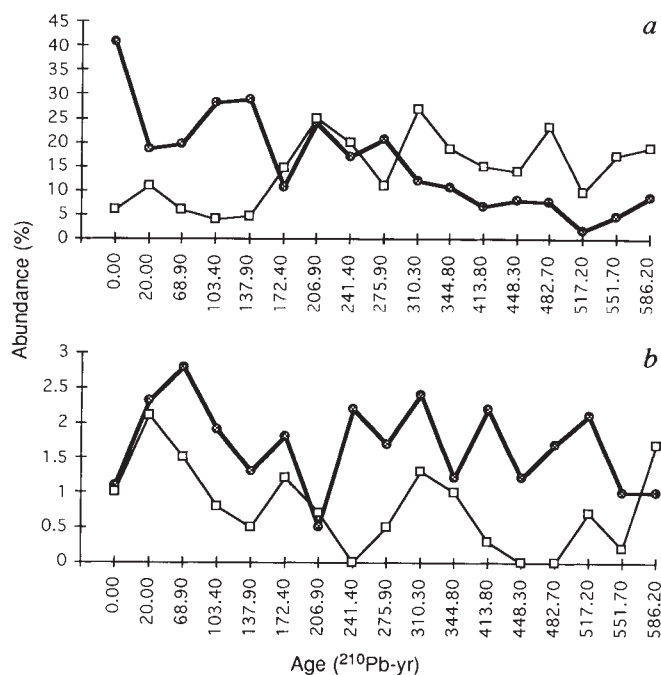


FIG. 2 Abundance of dominant phytoplankton and sea-ice taxa over the past 600 years in a sediment cores from Ellis Fjord in the Vestfold Hills. For each species, abundance over the past 20 years is within the range of abundances of the past 600 years. a, Long-term relative abundance fluctuations in *Chaetoceros* spp. vegetative cells (shaded circles, heavy lines) and *Fragilariopsis cylindrus* (empty squares, lighter lines) (ratio of *Chaetoceros* vegetative cells to resting cysts parallels rise in vegetative cells). b, Long-term relative abundance fluctuations in *Thalassiosira dichotomica* (empty squares, light lines) and *Entomoneis kjellmannii* (shaded circles, heavy lines). The first three of these species are significant elements of the summer phytoplankton bloom in Ellis Fjord, whereas *E. kjellmannii* dominates there in the spring sea-ice bloom¹⁴.

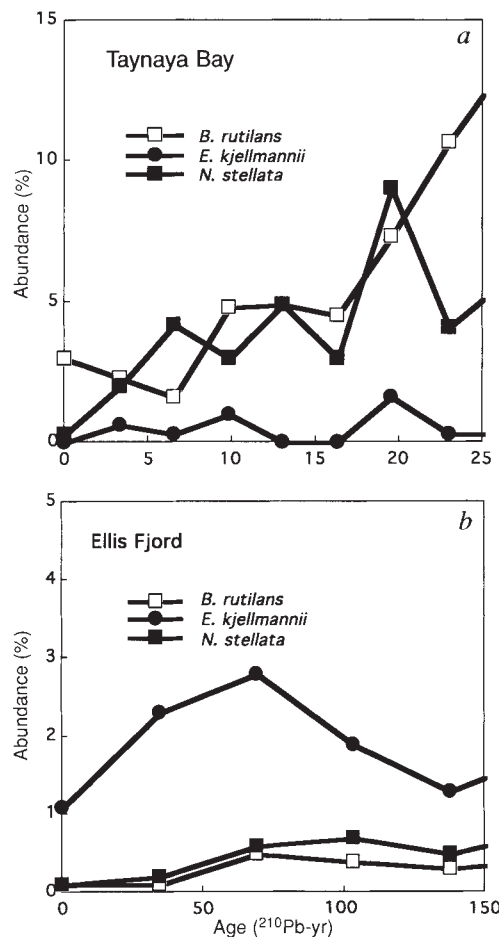


FIG. 3 Abundance of sea-ice taxa in cores from Taynaya Bay and Ellis Fjord. Abundances in the core from Long Fjord were too low to provide usable data. a, Relative abundance of sea-ice diatoms in Taynaya Bay. b, Relative abundance of sea-ice diatoms in Ellis Fjord. Data points from Taynaya Bay and Ellis Fjord were taken at 1-cm intervals.

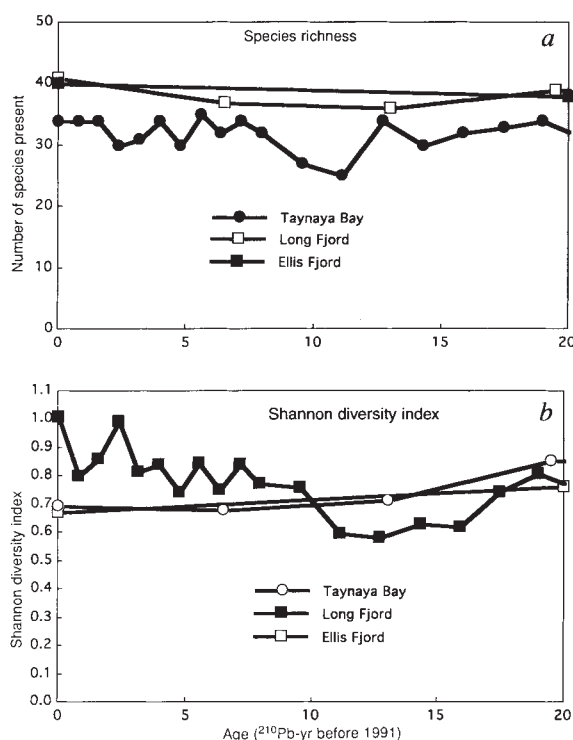


FIG. 4 a, Species richness over the past 20 years; data from Taynaya Bay, Long Fjord and Ellis Fjord sediment cores. b, Diversity (Shannon's diversity index) of diatom assemblages over the same time period and at the same locations.

the patchy spatial distribution of these communities⁸ may also contribute to their apparent temporal decline. This group of organisms is potentially the most threatened by an increase in ambient UVB but as yet there has been comparatively little investigation of their susceptibility. A recent report of a 5% reduction in primary productivity by some sea-ice algal species exposed to UVB in culture supports this suggestion⁹.

If increasing UVB irradiances were adversely affecting the diatom community then the number of species present (species richness) and the species diversity might be expected to decline. The assemblages from the fjords of the Vestfold Hills show little evidence of such changes over the past 20 years. The species richness in each core fluctuates within a relatively narrow range and no significant decline is indicated (Fig. 4a). Certainly there is no evidence that any species has become extinct in the fjords over the past 20 years. The Shannon diversity index likewise shows no appreciable variation over the examined time interval (Fig. 4b). These characteristics all suggest that increased UVB irradiances have had little effect on the planktonic diatom community.

Climate factors affect marine algal growth and competitiveness but there is no evidence over this time period for changes in average surface temperature, cloud cover, wind speed, water temperature or ice cover in the Vestfold Hills^{15,16}. There is no data available on variations in non-climate controls on productivity such as nutrient levels or predation, but these are unlikely to have changed significantly over this period. So, as there appears to have been no change in the diatom community, it is reasonable to suggest that the change in the UVB irradiance has been insufficient to influence diatom growth and competitiveness.

Evidence against the decline in abundance of sea-ice diatoms is equivocal as there is some suggestion of a decline in abundance of these taxa. Antarctic diatoms mostly have doubling rates of

2 to 5 days (ref. 17) so it is likely that any significant changes in abundance that could result from changes in UVB irradiances have already occurred, and that future changes in the relative composition of the diatom community are likely to be insignificant.

Minimal compositional changes in 20-year sequences of diatom assemblages from the Vestfold Hills suggest that the enhanced UVB levels resulting from the 'ozone hole' has had little effect on the diatom component of the phytoplankton community. In coastal areas such as this one, the presence of an extensive ice cover at the time of maximum ozone depletion and the timing of the major phytoplankton blooms after UVB levels have returned to pre-late-1970s long-term averages mitigate UVB-induced change. Our findings are not necessarily applicable to ice-edge and sea-ice communities where these conditions do not apply.

Received 23 February; accepted 1 July 1994.

1. El-Sayed, S. Z. *Nat. Hist.* **10**, 73–80 (1988).
2. Karentz, D. *Antarct. Sci.* **3**, 3–11 (1991).
3. Marchant, H. J., Davidson, A. T. & Kelly, G. M. *Biol.* **109**, 391–395 (1991).
4. El-Sayed, S. Z., Stephens, F. C., Bridgare, R. R. & Ondrusek, M. E. in *Antarctic Ecosystems. Ecological Change and Conservation* (eds Kerry, K. R. & Hempel, G.) 379–385 (Springer, Berlin, 1990).
5. Karentz, D., Cleaver, J. E. & Mitchell, D. L. *J. Phycol.* **27**, 326–341 (1991).
6. Davidson, A., Bramich, D., Marchant, H. & McMinn, A. *Mar. Biol.* (in the press).
7. Smith, W. O. & Nelson, D. M. *Science* **227**, 163–166 (1985).
8. Watanabe, K. *Jap. Soc. Phycol.* **36**, 221–229 (1988).
9. Palmisano, A. C. & Sullivan, C. W. in *Antarctic Nutrient Cycles and Food Webs* (eds Seigfried, W. R., Condy, P. R. & Laws, R. M.) 84–88 (Springer, Berlin, 1986).
10. Ryan, K. G. *J. Photochem. Photobiol.* **13**, 235–240 (1992).
11. Trodahl, H. J. & Buckley, R. G. *Science* **245**, 194–195 (1989).
12. Goldberg, E. D., Gambles, E., Griffin, J. J. & Koide, M. *Estuar. Coast. mar. Sci.* **5**, 549–561 (1977).
13. Robbins, J. A. & Edgington, D. N. *Geochim. cosmochim. Acta* **39**, 285–304 (1975).
14. McMinn, A. & Hodgson, D. J. *Plankton Res.* **15**, 925–938 (1993).
15. Jacka, T. H., Christou, L. & Cook, B. J. *ANARE Res. Notes* **22**, 1–97 (1984).
16. Russell-Head, D. & Simmonds, I. *Temporal Structure of Surface Weather Parameters at Casey, Davis, Mawson and Macquarie Island* (Publ. No. 35, School of Earth Sci., Univ. Melbourne, 1993).
17. Fiala, M. & Oriol, L. *Polar Biol.* **10**, 629–636 (1990).
18. Appleby, P. G. & Oldfield, F. in *Uranium Series Disequilibrium, Applications to Earth, Marine and Environmental Sciences* (eds Ivanovich, R. S. & Harmon, R. S.) (Clarendon, Oxford, 1992).

ACKNOWLEDGEMENTS. We thank the Australian National Antarctic Research Expeditions (ANARE) for logistical and field support, and G. Paltridge and P. Franzmann for reviewing the manuscript. This work was supported by the Antarctic Scientific Advisory Committee (ASAC) and the Australian Institute of Nuclear Science and Engineering (AINSE).

Sorptive preservation of labile organic matter in marine sediments

Richard G. Kell*, Daniel B. Montluçon*, Fredrick G. Prahl† & John I. Hedges*

* School of Oceanography, WB-10, University of Washington, Seattle, Washington 98195, USA

† College of Oceanography, Oregon State University, Corvallis, Oregon 97331, USA

ORGANIC matter preserved in marine sediments provides a molecular record of marine biological processes¹, accounts for approximately 20% of all carbon burial² and plays a key role in balancing the long-term flux of oxygen to the atmosphere³. Only recently has it been appreciated that more than 90% of the organic matter preserved in most marine sediments is intimately associated with mineral surfaces⁴. Little is known, however, of the effect that sorption to mineral surfaces might have in controlling either the lability or quantity of organic matter in the marine sedimentary record. The preserved organic material could be either intrinsically stable, or stabilized through interactions with mineral matrices. We show here that sorption of organic matter to mineral surfaces in marine

sediments stabilizes the component molecules, slowing remineralization rates by up to five orders of magnitude. Sorptive protection can therefore account for the enigmatic preservation of intrinsically labile molecules such as amino acids and simple sugars in marine deposits^{5,6} and links the preservation of organic carbon in marine sediments to the deposition of mineral surfaces.

Sediment was collected with a box corer from the upper continental slope of the Washington coast, USA ($46^{\circ}45.02'N$, $124^{\circ}59.87'W$; water depth 647 m) in August 1993. The core was kept at $2^{\circ}C$ until sectioned (~ 96 h) under N_2 into 0.5- (top 2 cm) or 1.0-cm slices. ^{210}Pb profiles of a sister core indicate that the sample had a mixed layer 5 cm deep, a stable sedimentation history throughout the first 20 cm and a sedimentation rate of $36\text{ mg cm}^{-2}\text{ yr}^{-1}$ (Table 1). Sediment within the core was dominated by minerals and composed largely of quartz, chlorite and smectite⁷, changing gradually from a well sorted surface sample (mean grain size $5\text{ }\mu\text{m}$) to poorly sorted silts deeper in the core (mean grain size $14.5\text{ }\mu\text{m}$). Changes in mineral grain size over the length of the core were reflected by a decrease in the mineral surface area from 45 to $20\text{ m}^2\text{ g}^{-1}$ (Fig. 1).

Organic carbon (OC) concentrations decreased downcore from 3.2% to $<1.5\%$. The 50% decrease in organic matter (OM) concentration covaried with a 50% decrease in mineral surface area (Fig. 1). Throughout the core, $>90\%$ of the OM could not be separated from the mineral matrix by either size or density fractionation (Table 1). The close association between OM and surface area suggests that OM loadings are controlled by the amount of mineral surface available for sorption. OC/surface area ratios in the core decreased slightly from 0.77 mg OC m^{-2} in modern sediment to 0.64 mg OC m^{-2} in the oldest (500 yr before present, BP) portion of the core (Fig. 1). These ratios fall in the range of $0.5\text{--}1.0\text{ mg OC m}^{-2}$ predicted by theory for monolayers of organic matter coating mineral surfaces^{4,8}, and overlap with ratios observed in other coastal sediments^{4,7,8} ($0.68\text{--}0.76\text{ mg OC m}^{-2}$). Thus within this core and other continental margin sediments^{4,7}, nearly all the OM is mineral-associated in a manner consistent with monolayer-equivalent sorption to mineral surfaces. Despite a 50% downcore decrease in OM loading, the OC/surface area ratio decreased by only 15%, suggesting that OM preservation was determined largely by changes in mineral surface area, and only marginally by *in situ* OM diagenesis.

Assuming that mineral surface area controls the quantity of OM preserved in these sediments, a question arises as to the extent to which mineral associations affect OM reactivity. The traditional view is that preserved sedimentary OM is intrinsically refractory⁹, in which case sorption to mineral surfaces is of little consequence. An alternate hypothesis is that sorption stabilizes intrinsically reactive OM. To test these hypotheses, four sections of the core (designated A–D) were used in partitioning and degradation experiments. The times since deposition (estimated by the ^{210}Pb activities) for the four depths in the core correspond to modern (<50), 160, 300 and 470 yr BP (Table 1). Reversibly sorbed organic matter was removed from the minerals by mild repetitive desorption using ultraviolet-oxidized sea water, 2N KCl and distilled water, each applied five times (Table 1). In general, distilled water removed 1.5 times more OM than either of the other treatments. Dimensionless partition coefficients (K) for the reversibly-sorbed OM in all samples and treatments averaged 0.08 (Table 1). These values are at the low end of the range ($0.07\text{--}8.0$) estimated for individual molecules such as amino acids and monosaccharides in marine sediments¹⁰. After the 15 desorption treatments, the amount of OM dissolved at each step had decreased from an average of 65 mg l^{-1} to $<2\text{ mg l}^{-1}$, indicating that nearly all the reversibly sorbed organic matter had been released: 25–50% of the original organic carbon was desorbed from the mineral matrix (Table 1).

After partitioning, the desorbed OM was exposed to free-living microorganisms from the overlying water column (Fig. 2). In all cases, $>70\%$ of the desorbed organic matter was remineral-

ized within 6 d by the growing bacterial assemblage (Fig. 2). The increase in bacterial abundance in the experimental containers, and the lack of a decrease in desorbed OM in control containers (no microorganisms present), indicate that changes in OM were not due to physical losses, but to bacterial utilization (Fig. 2, Table 1). OM concentrations in sediment slurries similarly incubated in the presence of oxygen changed by $<6\%$ (Fig. 2), suggesting that degradation was not due to enhanced efficiency of aerobic microorganisms relative to anaerobic microorganisms, but to liberation of the OM from mineral surfaces.

First-order degradation rate constants (k) for desorbed OM were $0.40 \pm 0.15\text{ d}^{-1}$ (Table 1). The youngest desorbed material (A and B) was remineralized most effectively, showing a 90% decrease in dissolved organic carbon, whereas OM from older sediments (C and D) was degraded to a lesser extent (Fig. 2). The microbial community also grew less efficiently on desorbed OM from deeper in the core. Of the total OM desorbed from young and older sediments, 30 and 10% of the utilized carbon was incorporated into biomass, respectively (Table 1). The decreasing reactivity of the reversibly-sorbed OM down the core may result from concentration of inherently refractory materials. Despite changes in the percentage of OM that could be desorbed, the amount of OM desorbed but not degraded (residual) remained fairly consistent at 0.13 mg g^{-1} throughout the core (Table 1). Thus, a small but constant component of the OM in the core resisted degradation.

Despite being present within the sediment for up to 500 yr, once desorbed from mineral surfaces, nearly all the OM was remineralized within days. Of the total sedimentary OM, 17–45%

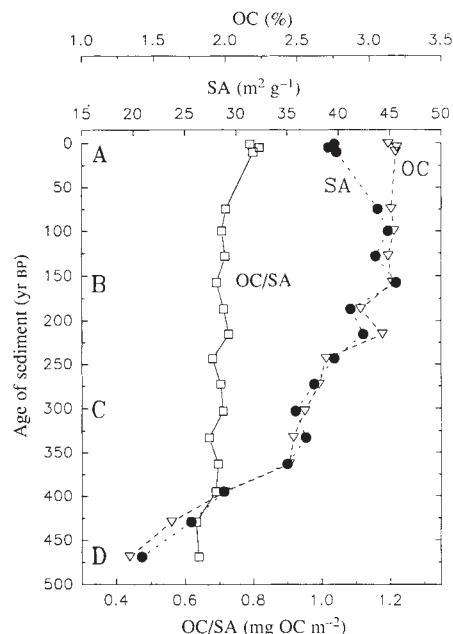


FIG. 1 Organic carbon (OC), mineral surface area (SA) and OC/SA ratio down a continental margin core from off the Washington coast, USA. Sub-samples used in the desorption and degradation experiment (Fig. 2) are listed as A–D, and were taken from the 1–2, 5–7, 19–24 and 32–36 cm sections of the core. OC was measured using a Carlo Erba CHN analyser with a precision of $\pm 0.25\%$ (ref. 7). Mineral surface area was determined by N_2 adsorption to organic-free mineral surfaces using the BET method, with a precision of $\pm 4\%$ (refs 7, 8). Average grain sizes for samples A–D were determined by standard pipette and settling techniques and correspond to 5, 8, 13.5 and $15.5\text{ }\mu\text{m}$, respectively. The age of the sediment was determined from ^{210}Pb dates, estimated from a sister core after accounting for differential compaction between the two cores using porosity profiles. Estimated error on the sediment age is ± 50 yr, due largely to bioturbation within the mixed layer.

TABLE 1 Data from marine sediments

	Age (yr BP)	Fraction of OM initially sorbed (%)	Fraction of OM desorbed (%)	K	(C/N) _i	(C/N) _f	(C/N) _d	GE (%)	Residual (mg g ⁻¹)	k (d ⁻¹)
A	<50	91 ± 5	50 ± 7	0.05	12.5	9.3	19.2	30 ± 4	0.15 ± 0.04	0.67 ± 0.05
B	160	96 ± 6	26 ± 6	0.08	11.6	10.5	12.0	24 ± 7	0.09 ± 0.06	0.45 ± 0.03
C	300	93 ± 5	23 ± 4	0.10	11.8	12.3	11.7	10 ± 5	0.14 ± 0.04	0.29 ± 0.03
D	470	95 ± 4	33 ± 9	0.07	13.2	11.7	14.1	12 ± 8	0.15 ± 0.02	0.21 ± 0.05

Fraction of organic matter (OM) initially sorbed to mineral surfaces was estimated by density fractionation using CsCl (density $\rho = 1.9 \text{ g cm}^{-3}$)^{7,8}. Fraction of OM desorbed was determined from the mass balance between OM loss from the minerals and the increase in dissolved OM during desorption treatments. Desorbed OM measurements agreed with losses in the sedimentary OM to within $\pm 5\%$. Dimensionless partition coefficients (K) (ref. 10) were estimated from measurements of both sedimentary OM and desorbed OM (measured with a Shimadzu TOC-5000 dissolved-carbon analyser). For desorption treatments, a 1 g sediment: 4 ml suspension was used to repetitively partition OM into ultraviolet-oxidized sea water, 2N KCl, and distilled water (5 times each). Samples were constantly shaken and held at 4 °C. After 2–12 h each treatment was centrifuged and filtered (0.2 μm), and the fraction of OM desorbed into the liquid phase was measured. (C/N)_i, initial atomic carbon/nitrogen ratio of the OM; (C/N)_f, final ratio for the irreversibly sorbed OM; (C/N)_d, ratio for the desorbed OM; GE, growth efficiency of the bacterial assemblage estimated as the fraction of desorbed OM incorporated into bacterial biomass divided by the total decrease in desorbed OM, $\times 100$ (ref. 18); Residual, amount of OM per gram of sediment that was desorbed but not degraded; k , first-order rate constant for degradation.

was remineralized. Other studies have shown similar stabilization effects for monomeric compounds such as amino acids^{10–12}, but polymers such as proteins can be either preserved or degraded at enhanced rates¹³ when sorbed to mineral surfaces. Many organic pollutants are also subject to either enhanced or inhibited degradation when sorbed to minerals^{14,15}. Published data on sorptive effects are equivocal, due largely to the wide

array of experimental conditions employed¹⁵. Our study of bulk OM reactivity within marine sediments is the first to use total OM, rather than specific compounds, to address the question of stabilization effects, and clearly shows that a large portion of the OM present in marine sediments is intrinsically reactive material stabilized only because it is within a sedimentary milieu.

The degradation rate of desorbed OM can be compared to the *in situ* degradation rate. We estimated *in situ* degradation from the decrease in the OC/surface area ratio. The resulting first-order rate constant for OM degradation within the 500-yr core section was $7.7 \times 10^{-7} \text{ d}^{-1}$. This is 5×10^5 times less than the degradation rate of desorbed OM ($0.40 \pm 0.15 \text{ d}^{-1}$). If the sorptively-driven relationship between OM and surface area is ignored, and OM diagenesis is modelled from decreases in percentage OM, the calculated *in situ* degradation rate constant rises to only $5.1 \times 10^{-6} \text{ d}^{-1}$. Regardless of the method used to estimate *in situ* degradation, desorbed OM was degraded at a rate that is five orders of magnitude faster.

Control of OM preservation by mineral surface area indicates that it is the availability of surface area, not the supply of OM, that limits the extent of OM preservation^{4,7,8}. Sorptive protection thus decouples preservation from primary productivity and links it to the delivery of mineral grains from eroding continents. This provides a mechanistic foundation for the strong empirical relationships between mineral sedimentation rate and OM concentration¹⁶ or burial efficiency⁶. Also, sorptive protection provides a preservation mechanism that goes hand-in-hand with kerogen weathering on continents. A balance between these two physically separate processes has been traditionally assumed in models of global O₂ and CO₂ cycles^{2,3,17}, but has been difficult to explain. If burial of OM were controlled by productivity or other extrinsic factors, there would be no *a priori* reason for OM burial to balance kerogen weathering. But if weathering and burial are both controlled by the erosion (and deposition) of sedimentary rocks, the link is established. As most continental margin sediments contain monolayer-equivalent amounts of OM^{7,8}, and as these sediments are the site of nearly all modern OM preservation^{2,5–7,17}, stabilization of OM by interactions with mineral matrices may be the single largest factor controlling OM preservation on the Earth's surface today. □

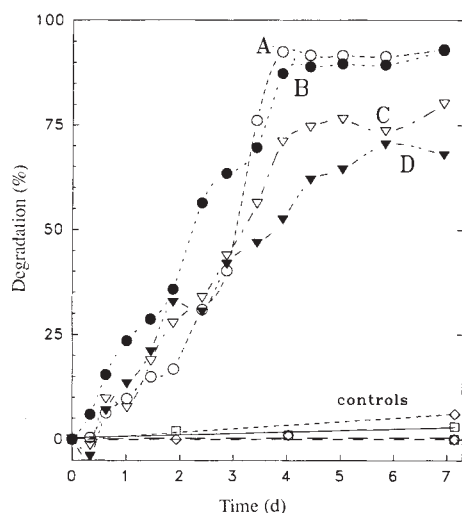


FIG. 2 Degradation of desorbed organic matter (OM) from four depths (A–D) in the studied sediment core (Table 1) by natural assemblages of aerobic microorganisms. After repetitive desorption with organic-free solutions (blanks were $<2\%$ of the amount of OM desorbed), all 15 solutions containing released organic matter were combined and further diluted with ultraviolet-oxidized sea water (organic-free) which resulted in a dissolved organic matter pool of $5\text{--}7 \text{ mg C l}^{-1}$. A sample (10 ml) of whole sea water containing the natural microbial assemblage was collected from surface waters above the sampling location, and added to 1 l of desorbed OM solution. OM was then monitored over time (both 0.2- μm filtered water and on unfiltered water⁷), as was the growth of the microbial assemblage¹⁸. Percentage degradation was estimated as 100 minus the percentage change in dissolved OM ($<0.2 \mu\text{m}$; that is, degradation is remineralization plus production of filterable biomass) over time. Control experiments (□) were conducted without additions of the microbial assemblage in order to account for sorption of the organic matter to the walls of the glass container. These consisted of 6.1 and 7.0 mg l^{-1} additions of dissolved organic carbon from samples A and C, respectively. Slurry experiments (◇) consist of 1:2 mixtures (w/v) of whole sea water and unaltered sediment which were allowed to incubate on a shaker table at $\sim 22^\circ \text{C}$ for 431 h.

Received 28 March; accepted 14 July 1994.

- Summons, R. E. in *Organic Geochemistry* (eds Engel, M. & Macko, S.) 1–21 (Plenum, New York, 1993).
- Berner, R. A. *Am. J. Sci.* **282**, 451–473 (1982).
- Berner, R. A. *Glob. planet. Change* **1**, 97–122 (1989).
- Mayer, L. M. in *Organic Geochemistry* (eds Engel, M. & Macko, S.) 171–183 (Plenum, New York, 1993).
- Hedges, J. I. *Mar. Chem.* **39**, 67–94 (1992).
- Henrichs, S. M. in *Organic Geochemistry* (eds Engel, M. & Macko, S.) 101–117 (Plenum, New York, 1993).

7. Keil, R. G., Tsamakis, E., Fuh, C. B., Giddings, J. C. & Hedges, J. I. *Geochim. cosmochim. Acta* **58**, 879–893 (1994).
8. Mayer, L. M. *Geochim. cosmochim. Acta* **58**, 1271–1284 (1994).
9. Eglinton, G. in *Organic Geochemistry* (eds Eglinton, G. & Murphy, M. T. J.) 20–71 (Springer, New York, 1969).
10. Henrichs, S. M. & Sugai, S. F. *Geochim. cosmochim. Acta* **57**, 823–835 (1993).
11. Pinck, L. A. & Allison, F. E. *Science* **114**, 130–131 (1951).
12. Wang, X.-C. & Lee, C. *Mar. Chem.* **44**, 1–24 (1993).
13. Marshman, N. A. & Marshall, K. C. *Soil Biol. Biochem.* **13**, 127–134 (1981).
14. Ogram, A. V. et al. *Appl. envir. Microbiol.* **49**, 582–587 (1985).
15. van Loosdrecht, M. C. M. et al. *Microbiol. Rev.* **54**, 75–87 (1990).
16. Doyle, L. J. & Garrels, R. M. *Geo-Mar. Lett.* **5**, 51–53 (1985).
17. Berner, R. A. *Am. J. Sci.* **291**, 339–376 (1991).
18. Keil, R. G. & Kirchman, D. L. *Mar. Ecol. Prog. Ser.* **73**, 1–10 (1991).

ACKNOWLEDGEMENTS. We appreciate comments provided by R. Benner, M. Peterson, B. Bergamaschi and the UW Marine Organic Geochemistry group; L. Mayer provided encouragement and stimulating discussions. We thank E. Tsamakis, K. Rossner and C. Cooper for technical support, and Chih-An Huh from OSU for graciously providing the ^{210}Pb data. This research was supported by the US NSF.

Archaean crustal development in the Lewisian complex of northwest Scotland

Kevin W. Burton*, Anthony S. Cohen*,
R. Keith O'Nions* & Michael J. O'Hara†

* Department of Earth Sciences, University of Cambridge,
Downing Street, Cambridge CB2 3EQ, UK

† Department of Geology, University of Cardiff, Cardiff CF1 3YG, UK

THE Lewisian complex of northwest Scotland is typical of many Archaean terrains and has a well documented history starting ~2,700 Myr ago^{1–6}. Here we present new isotopic data that extend this history back to 3,300 Myr, and provide some insight into how the earliest continental crust may have formed. The Lewisian is dominated by tonalite, trondhjemite and granodiorite (TTG) rocks, which require a mafic lithospheric source, rather than being direct mantle melts^{7–11}. But the mafic and ultramafic rocks in the high-grade granulite-facies part of this terrain show little evidence of a significantly older crustal history^{12,13} and precursor material to the TTG lithologies has not yet been identified. Here we show that older amphibolite material has survived at a lower metamorphic grade. Coexisting amphibolite minerals yield indistinguishable ^{207}Pb – ^{206}Pb and ^{147}Sm – ^{143}Nd ages of $3,310 \pm 27$ Myr and $3,298 \pm 73$ Myr, respectively. These data are consistent with an origin for much of the Lewisian terrain by the re-melting of pre-existing lithosphere, with an isotopic signature similar to that of the amphibolites studied here.

Low-grade amphibolite-facies rocks in the Lewisian terrain occur at Gruinard Bay. Layered amphibolites and tonalites at this locality are intruded by trondhjemitic gneiss, forming a large-scale agmatite complex^{14,15}. Field relationships suggest that these amphibolites are one of the earliest parts of the Lewisian complex, and trace-element chemistry suggests that they are a potential source for many of the TTG rock types^{14,15}. Thus it is possible that they preserve valuable information on crustal evolution in this terrain. We have obtained samarium–neodymium and lead isotopic data for the amphibolite and trondhjemite mineral assemblages from Gruinard Bay, and compare our results with those of earlier studies of TTG lithologies in this terrain^{1–6}.

Chemical and isotopic data for mineral and whole-rock separates taken from both trondhjemite and amphibolite samples are given in Table 1 (see legend for details of mineral petrography). Analytical procedures follow those reported previously¹⁶. For the trondhjemites both ^{207}Pb – ^{206}Pb and ^{147}Sm – ^{143}Nd ages are in good agreement and indicate that equilibration of Sm–Nd and Pb between minerals occurred ~2,400 Myr ago

(Table 2). In contrast, Pb isotopic data for whole-rock and mineral separates from the amphibolite samples yields ^{207}Pb – ^{206}Pb versus ^{206}Pb – ^{204}Pb regression ages of $3,310 \pm 26$ Myr and $3,314 \pm 61$ Myr, respectively (Fig. 1a; Table 2). Sm–Nd data for sample GR102 give an age of $3,298 \pm 73$ Myr (Fig. 1b; Table 2). The Sm–Nd results for sample GR11 yield no meaningful age information, and these data are most probably biased by the presence of retrograde phases. Both ^{207}Pb – ^{206}Pb ages, and the ^{147}Sm – ^{143}Nd age for sample GR102, indicate that the mineral equilibration in the amphibolites occurred ~3,300 Myr ago. This

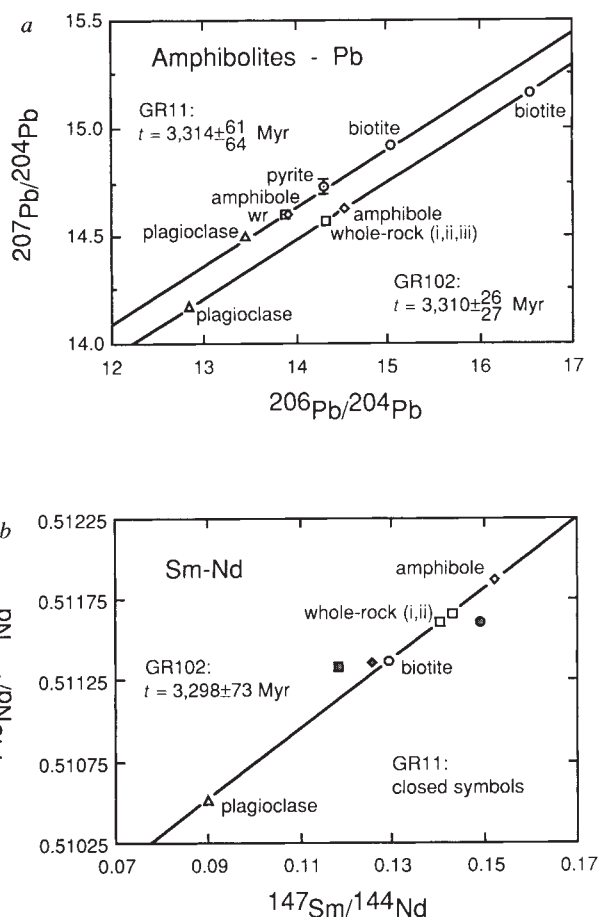


FIG. 1 a, ^{207}Pb – ^{206}Pb evolution diagram for amphibole (diamonds), biotite (circles), plagioclase (triangles), pyrite (dotted circle) and whole-rock (squares) separates from the amphibolite samples (GR102 and GR11) from Gruinard Bay. The data for sample GR102 yield a best-fit line corresponding to an age of $3,310 \pm 26$ Myr; mean-square weighted deviation is 0.78. Sample GR11 gives an age of $3,314 \pm 61$ Myr; mean-square weighted deviation is 4.25. b, ^{147}Sm – ^{143}Nd evolution diagram for the minerals shown in Fig. 2a. The data for sample GR102 yield a best-fit line corresponding to an age of $3,298 \pm 73$ Myr; mean-square weighted deviation is 0.27. Data from sample GR11 (filled symbols) are also shown, but do not yield any meaningful age information. The age obtained for sample GR102 is in excellent agreement with the ^{207}Pb – ^{206}Pb data from both amphibolite samples, and indicate that equilibration of both Sm–Nd and Pb between minerals in the amphibolites occurred 3,300 Myr ago. This age is ~600 Myr older than the time of depletion in the higher-grade granulites^{1–6}, and 900 Myr older than the minerals in the adjacent trondhjemites. (Note that the statistical precision of all of the isochrons both here and in Table 2 is excellent; however, all of the phases analysed have low U/Pb and Sm/Nd ratios. Under these circumstances the relatively large uncertainties in the age estimates arise as a result of the small amount of ingrowth of radiogenic Pb and Nd, rather than owing to analytical errors or geological disturbance.)

TABLE 1 Isotope data for trondhjemites and amphibolites at Gruinard Bay

Whole-rock/ mineral	$^{206}\text{Pb}/^{204}\text{Pb}^*$	$^{207}\text{Pb}/^{204}\text{Pb}^*$	$^{208}\text{Pb}/^{204}\text{Pb}^*$	$^{143}\text{Nd}/^{144}\text{Nd}^\dagger$	$^{147}\text{Sm}/^{144}\text{Nd}^\ddagger$	[Pb] (p.p.m. by wt)	[Sm] (p.p.m. by wt)	[Nd] (p.p.m. by wt)
Trondhjemites								
GR100								
Whole-rock	14.187 $\pm$ 0.014	14.595 $\pm$ 0.016	34.326 $\pm$ 0.038	0.511102 $\pm$ 26	0.1124	8.831	0.7261	3.904
Amphibole	15.791 $\pm$ 0.014	14.865 $\pm$ 0.016	38.242 $\pm$ 0.032	0.511777 $\pm$ 20	0.1548	2.208	1.353	5.284
Biotite	16.623 $\pm$ 0.016	14.993 $\pm$ 0.018	37.542 $\pm$ 0.034	0.511246 $\pm$ 20	0.1201	2.770	0.6087	3.063
Plagioclase	13.910 $\pm$ 0.010	14.578 $\pm$ 0.016	33.969 $\pm$ 0.028	0.511000 $\pm$ 26	0.1055	7.286	0.2972	1.703
GR103								
Whole-rock	13.628 $\pm$ 0.012	14.525 $\pm$ 0.016	35.559 $\pm$ 0.034	0.511170 $\pm$ 18	0.1138	9.508	4.868	25.86
Amphibole	14.561 $\pm$ 0.012	14.679 $\pm$ 0.016	36.933 $\pm$ 0.032	0.511871 $\pm$ 14	0.1591	2.757	9.720	36.92
Plagioclase	13.489 $\pm$ 0.010	14.496 $\pm$ 0.010	34.686 $\pm$ 0.026	0.510635 $\pm$ 20	0.0802	7.892	0.4862	3.664
Amphibolites								
GR102								
Whole-rock i§	14.321 $\pm$ 0.016	14.573 $\pm$ 0.016	37.096 $\pm$ 0.038	0.511604 $\pm$ 20	0.1401	1.930	11.58	49.97
Whole-rock ii§	14.324 $\pm$ 0.018	14.564 $\pm$ 0.018	36.929 $\pm$ 0.044	0.511661 $\pm$ 18	0.1430	1.924	11.77	49.75
Whole-rock iii	14.317 $\pm$ 0.012	14.566 $\pm$ 0.018	36.988 $\pm$ 0.036	—	—	1.578	—	—
Amphibole	14.544 $\pm$ 0.016	14.624 $\pm$ 0.016	36.796 $\pm$ 0.044	0.511873 $\pm$ 14	0.1524	2.551	17.07	67.72
Biotite	16.536 $\pm$ 0.016	15.172 $\pm$ 0.016	37.440 $\pm$ 0.038	0.511374 $\pm$ 20	0.1293	1.303	3.880	18.13
Plagioclase	12.855 $\pm$ 0.010	14.177 $\pm$ 0.016	35.220 $\pm$ 0.028	0.510511 $\pm$ 26	0.0901	2.706	1.341	8.991
GR11								
Whole-rock	13.897 $\pm$ 0.010	14.594 $\pm$ 0.016	37.177 $\pm$ 0.038	0.511330 $\pm$ 6	0.1182	1.221	4.927	25.18
Amphibole	13.912 $\pm$ 0.010	14.594 $\pm$ 0.016	36.683 $\pm$ 0.028	0.511360 $\pm$ 6	0.1259	1.714	14.04	67.37
Biotite	15.047 $\pm$ 0.010	14.919 $\pm$ 0.016	36.083 $\pm$ 0.028	0.511613 $\pm$ 8	0.1490	0.9732	1.336	5.419
Plagioclase	13.462 $\pm$ 0.010	14.498 $\pm$ 0.016	34.884 $\pm$ 0.028	—	—	2.339	—	—
Pyrite	14.297 $\pm$ 0.036	14.730 $\pm$ 0.036	36.241 $\pm$ 0.094	—	—	—	—	—

All quoted errors are $2\sigma_m$. Maximum procedural blanks are 60 pg for Pb, 5 pg for Nd and 2 pg for Sm.

All samples were taken from a large roadcut on the west side of Gruinard Bay (grid reference: NG941902). The trondhjemite mineral assemblage comprises plagioclase(An_{25})–quartz–hornblende–biotite–epidote–apatite–sphene. The amphibolite mineral assemblage consists of hornblende–biotite–plagioclase(An_{19})–calcite–clinozoisite–ilmenite($X_{\text{ilm}} = 0.943$)–magnetite($X_{\text{usp}} = 0.006$)–pyrite. (Electron microprobe analyses for the major mineral phases are available on request from K.W.B.) In the amphibolites studied here the original high-Ti pargasitic hornblende has been pseudomorphed by edenite with exsolved ilmenite; more rarely, hornblende shows sieve-textured quartz inclusions considered to be retrograde after clinopyroxene. Amphibolite GR11 shows extensive alteration and contains calcite, scapolite, alkali feldspar, chlorite and Fe–Ti oxides rimmed by sphene. The trondhjemites also show evidence of retrogression, with hornblende after clinopyroxene. Plagioclase in both rock types is sometimes altered to sericite, clinozoisite and muscovite.

* Pb isotope ratios relative to NBS981 standard, which yields $^{206}\text{Pb}/^{204}\text{Pb} = 16.932 \pm 0.012$, $^{207}\text{Pb}/^{204}\text{Pb} = 15.486 \pm 0.016$, $^{208}\text{Pb}/^{204}\text{Pb} = 36.691 \pm 0.036$ (2σ ; $n = 35$).

† $^{143}\text{Nd}/^{144}\text{Nd}$ normalized to $^{146}\text{Sm}/^{144}\text{Nd} = 0.7219$; JM-Nd standard yielded 0.511129 ± 18 (2σ , external; $n = 42$) for the period January 1987 to March 1991 (samples GR100, GR102, GR103) and 0.511125 ± 10 (2σ , external; $n = 26$) for the period January 1992 to October 1993 (sample GR11). By comparison, a single analysis of the La Jolla Nd standard yields 0.511856 ± 6 and no correction for interlaboratory bias is considered necessary.

‡ $^{147}\text{Sm}/^{144}\text{Nd}$ ratio determined to a precision of $\pm 0.1\%$.

§ Whole-rock powders repeated through dissolution, chemical separation and mass spectrometric analysis.

|| Uranium concentration determined for GR102 whole-rock iii; $[U] = 0.080$ p.p.m., $^{238}\text{U}/^{204}\text{Pb} = 2.93$.

age is ~ 600 Myr older than the time of granulite metamorphism in the higher-grade part of the terrain^{1–6} and 900 Myr older than the minerals in the adjacent trondhjemites (see below).

The initial ϵ_{Nd} value defined by the amphibolite minerals at 3,300 Myr (see Table 2) is close to the value expected for the depleted mantle at that time (Fig. 2a). This suggests that these rocks experienced metamorphism soon after differentiation from a depleted mantle source, and were unaffected by the later granulite-facies event. By comparison, a variety of TTG lithologies in the high-grade part of this terrain experienced Sm and Nd fractionation during granulite-facies metamorphism $\sim 2,700$ Myr ago, and preserve an initial ϵ_{Nd} value of -1.89 ± 1.58 (combined data from refs 4, 13 and 17). The trondhjemites at Gruinard Bay define initial ϵ_{Nd} values much lower than expected for the contemporaneous depleted mantle (Table 2), which points to a crustal source for these rocks. The evolution of the initial Nd isotope compositions from the amphibolites to the tonalites, and then to the trondhjemites, requires a $^{147}\text{Sm}/^{144}\text{Nd}$ ratio of 0.132 (the 'best-fit' regression shown as a dashed line in Fig. 2a)—within the observed range for continental crust^{18,19}. Thus, these data are consistent with an origin for the tonalites, and ultimately the trondhjemites, by partial melting of the older amphibolites, or similar material, formed as crust at 3,300 Myr ago. For the trondhjemites at Gruinard Bay this interpretation requires that their intrusion occurred soon before mineral equi-

libration 2,400 Myr ago. Whole-rock data for the trondhjemites give age estimates of between 2,800 and 3,000 Myr (refs 5, 12), which could be taken to indicate that intrusion occurred around that time. But if the ages reported here were to reflect final isotopic closure on the scale of the minerals, this would require that these minerals experienced complete diffusional re-equilibration of Pb and Nd isotopes, whereas the adjacent amphibolites containing similar minerals did not. Thus, the whole-rock ages are considered simply to reflect the age of the mantle source, in much the same way that the Scourie dykes have whole-rock ages of $\sim 3,000$ Myr (ref. 20), whereas their intrusion and crystallization clearly postdates granulite formation^{20–22}. Recent data suggest that this terrain experienced a second episode of high-grade metamorphism, and igneous intrusion, $\sim 2,490$ Myr ago⁶. Metamorphism and trondhjemite emplacement at Gruinard Bay may well have occurred at this time.

In contrast to the tonalites and trondhjemites just discussed, amphibolite-facies granodiorites from the northern part of the complex¹² cannot have evolved in a simple manner from material having an isotopic composition like that of the amphibolites at Gruinard Bay. However, it is difficult to assess the significance of these data because of the large uncertainties associated with these results^{4,5,12}. Moreover, some of these rocks have experienced uranium depletion⁵, and the same samples have unusually low Sm/Nd ratios. If the fractionation of U/Pb and Sm/Nd

TABLE 2 Mineral ages and initial isotopic compositions

	Age (Myr)		$^{143}\text{Nd}/^{144}\text{Nd}_i$	$^{206}\text{Pb}/^{204}\text{Pb}_i$	μ_1	μ_{WR}	ϵ_{Nd}^i
	$^{206}\text{Pb}/^{207}\text{Pb}$	$^{147}\text{Sm}/^{143}\text{Nd}$					
Trondhjemites							
GR100	$2,431 \pm 68$	$2,385 \pm 87$	0.509346 ± 42	13.530 ± 0.088	7.38 ± 0.15	1.43 ± 0.22	-3.90 ± 0.82
GR103	$2,544 \pm 179$	$2,373 \pm 51$	0.509384 ± 28	13.382 ± 0.204	7.46 ± 0.37	0.51 ± 0.45	-3.44 ± 0.55
Amphibolites							
GR102	$3,310 \pm 26$	$3,298 \pm 73$	0.508549 ± 31	12.186 ± 0.030	8.02 ± 0.08	3.18 ± 0.07	$+3.89 \pm 0.61$
GR11	$3,314 \pm 64$	—	—	$12,326 \pm 0.074$	8.43 ± 0.21	2.43 ± 0.13	—

All quoted errors are $\pm 2\sigma$. Definitions of quantities: $\mu_1 = ({}^{238}\text{U}/{}^{204}\text{Pb}) = \{({}^{206}\text{Pb}/{}^{204}\text{Pb})_i - \{({}^{206}\text{Pb}/{}^{204}\text{Pb})_i\} / \exp(\lambda_{238}T - \exp(\lambda_{238}t))\}$; $\mu_{\text{WR}} = ({}^{238}\text{U}/{}^{204}\text{Pb})_{\text{WR}} = \{({}^{206}\text{Pb}/{}^{204}\text{Pb})_m - \{({}^{206}\text{Pb}/{}^{204}\text{Pb})_i\} / (\exp(\lambda_{238}t) - 1)\}$; $\epsilon_{\text{Nd}} = [({}^{143}\text{Nd}/{}^{144}\text{Nd})_{\text{sample}} / ({}^{143}\text{Nd}/{}^{144}\text{Nd})_{\text{CHUR}} - 1] \times 10^4$; $f_{\text{Sm}/\text{Nd}} = \{({}^{147}\text{Sm}/{}^{144}\text{Nd})_{\text{sample}} / ({}^{147}\text{Sm}/{}^{144}\text{Nd})_{\text{CHUR}}\} - 1$. Where T is the age of the Earth, 4,565 Myr (ref. 28); t is the ${}^{207}\text{Pb}$ - ${}^{206}\text{Pb}$ age (given in column 1); λ_{238} is the decay constant for ${}^{238}\text{U}$ (ref. 29). Subscripts 'i', 't' and 'm' indicate the initial Pb ratios at 4,565 Myr (ref. 30); the Pb ratios at time t (given in column 4) and the present-day measured Pb ratios for the whole-rock samples reported in Table 1, respectively. $({}^{143}\text{Nd}/{}^{144}\text{Nd})_{\text{sample},i}$ is the initial ${}^{143}\text{Nd}/{}^{144}\text{Nd}$ ratio defined by the best-fit regression line for each mineral dataset at time t (given in column 3). $({}^{143}\text{Nd}/{}^{144}\text{Nd})_{\text{CHUR},i}$ is the ${}^{143}\text{Nd}/{}^{144}\text{Nd}$ ratio of the chondrite uniform reservoir (CHUR)³¹ at time t ; error propagation follows the method of Fletcher and Rosman³².

occurred after their formation, then any age or isotopic information should be interpreted with extreme caution.

The Pb isotope data can also be used to place some constraints on the behaviour of U and Pb in these rocks. The amphibolite data give model μ_1 values (the ${}^{238}\text{U}/{}^{204}\text{Pb}$ ratio which describes evolution in the source rocks from 4,565 Myr to the determined age, see Table 2) close to recent estimates for the depleted mantle^{23,24}. A second μ value may be calculated from the whole-rock data: this model μ_{WR} reflects the ${}^{238}\text{U}/{}^{204}\text{Pb}$ ratio for the sample from its formation time to the present day. Values of μ_{WR} for the trondhjemites are low (Table 2), which may either be attributed to uranium depletion during metamorphism, or may reflect the low μ of their source. However, the μ_{WR} values of 3.18 and 2.43 for the amphibolites are also low (Table 2), in good agreement with the present day ${}^{238}\text{U}/{}^{204}\text{Pb}$ ratio of 2.93

measured for sample GR102 (see Table 1 legend). This may indicate that the amphibolites also experienced a reduction in their U/Pb ratio during metamorphism 3,300 Myr ago. Alternatively, this may be due to the magmatic processes responsible for the generation of the precursors of the amphibolites, during which the source μ of ~ 8.2 was reduced to a value of ~ 3 . In either case, if the tonalites and the trondhjemites were indeed produced by partial melting of this older amphibolite crust then their μ_1 values should preserve some evidence of a crustal pre-history, from 3,300 Myr to their determined age, with a low μ of ~ 3 . Figure 2b indicates that the μ_1 values for the amphibolites, tonalites and trondhjemites with ages ranging from 3,300 to 2,400 Myr may be related. If it is assumed that the dashed line shown in Fig. 2b reflects the Pb evolution in a lithospheric source over this time interval, then the particular μ value for that source

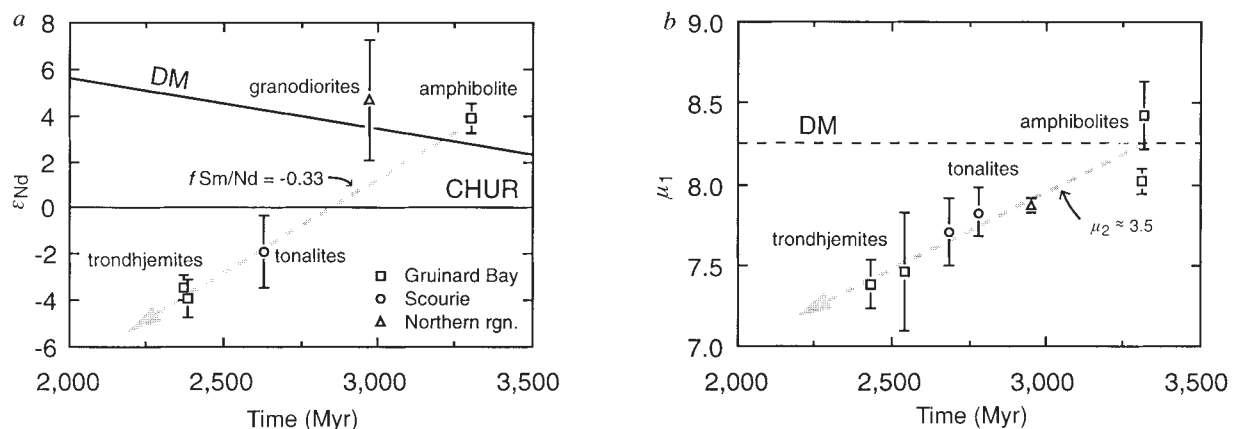


FIG. 2 a, Plot of initial ϵ_{Nd} against time (Myr ago) for the amphibolite and trondhjemite minerals from Gruinard Bay (squares), for the granulite-facies TTG lithologies (circles) at Scourie^{4,13,17}, and amphibolite-facies granodiorites (triangle) from the northern part of the Lewisian complex¹². All data are expressed relative to the chondrite uniform reservoir (CHUR)³¹, defined as $\epsilon_{\text{Nd}} = 0$. Also shown is the model depleted-mantle (DM) evolution curve of Goldstein *et al.*¹⁹. The initial ϵ_{Nd} value for the amphibolite minerals is close to that expected for the depleted mantle at that time (see text). The apparent evolution of ϵ_{Nd} values from the amphibolites to the tonalites, and then to the trondhjemites (the best-fit regression shown as a dashed line) requires an $f_{\text{Sm}/\text{Nd}}$ value of -0.33 (where $f_{\text{Sm}/\text{Nd}}$ is the Sm/Nd ratio relative to CHUR; see Table 2 legend)—within the range of values for the continental crust^{18,19}. Note, however, that the amphibolite-facies granodiorites from the northern region do not fall on the line, and therefore cannot have evolved in this simple manner (see text). These data suggest that many of the TTG lithologies in the Lewisian may have been formed by partial melting of the older amphibolite crust, or similar material, preserved at Gruinard Bay. b, Comparison of μ_1 values (${}^{238}\text{U}/{}^{204}\text{Pb}$ ratios which describe Pb isotope evolution from 4,565 Myr to time t) against time (Myr ago) for

the amphibolite and trondhjemite minerals from Gruinard Bay, and the whole-rock tonalite data from Scourie and the northern region (data from refs 3–5). The amphibolites from Gruinard Bay give μ_1 values close to a recent estimate of 8.26 ± 0.1 for the depleted mantle²⁴, shown as a dashed line (labelled DM). At 3,300 Myr the source μ of the amphibolites was reduced to a value of ~ 3 (see Table 2). If the variation of μ_1 values with time in the TTG lithologies is related to the development of a single lithospheric source region, then its characteristic ${}^{238}\text{U}/{}^{204}\text{Pb}$ (μ_2) value for the time interval 3,300 to 2,400 Myr is defined from the slope of the best-fit regression, shown as a dashed line. The best estimate is $\mu_2 \approx 3.5$. Thus, these data suggest that the TTG lithologies may have been derived from a lithospheric source with a low μ , similar to that of the amphibolites. (Definitions of quantities: $\mu_2 = (\mu^* (\exp(\lambda_{238}T) - 1)) / ((\mu_1 (\exp(\lambda_{238}T) - 1)) / (\exp(\lambda_{238}t_1) - 1))$, where μ_1 is the μ value from 4,565 Myr to t_1 (3,300 Myr), μ^* is the μ value from 4,565 Myr to t_2 (the time of Pb isotope equilibration, μ_2 is the μ value from t_1 (3,300 Myr) to t_2 , T is the age of the Earth, 4,565 Myr (ref. 28) and λ_{238} is the decay constant for ${}^{238}\text{U}$ (ref. 29). From the best-fit line shown when $t = 3,300$ Myr, $\mu_1 = \mu^* = 8.23$; when $t = 0$, $\mu^* = \text{intercept} = 5.13$.)

is calculated to be $\mu_2 \approx 3.5$ (see Fig. 2b legend). This result gives no constraint on the time of formation of the TTG rock types, but simply indicates that they may have been derived from a lithospheric source with a low μ , similar to that of the amphibolites at Gruinard Bay. Many granulites show extreme depletion in uranium^{25,26}, which gives rise to very low μ values. But in the present case the low μ values must have developed, at least in part, before granulite-facies metamorphism and are related either to magmatic processes²⁷ or to lower-grade amphibolite-facies metamorphism¹ rather than to granulite formation itself.

The Lewisian complex, like many Archaean terrains and hence the bulk of the early continental crust, is dominated by TTG rock types. Consequently, understanding the melting processes responsible for their generation, and the timescale over which this occurs, are central to ideas on early continent formation. Petrogenic and trace-element constraints suggest that these rocks are formed by partial melting of pre-existing mafic crust⁷⁻¹¹. But before this work, potential source material had eluded detection and the Lewisian TTG rock types were considered to have formed within ~300 Myr of granulite metamorphism at ~2,700 Myr. Our results indicate that the amphibolites at Gruinard Bay have been in existence for at least 3,300 Myr, and thus the Lewisian TTG rocks may have developed over a much longer time interval than previously assumed. The amphibolites possess initial isotopic compositions close to those anticipated for the depleted mantle at that time. Furthermore, the apparent evolution of Nd and Pb isotopes within many of the TTG lithologies in the Lewisian suggests that they may have been derived from a pre-existing lithospheric source, with an isotopic signature similar to that of the amphibolites at Gruinard Bay. It seems possible that multiple stages of lithospheric processing, such as described here, may characterize the development of TTG rocks, both in the Lewisian complex and in other Archaean terrains. Under these circumstances only the earliest remnants in such terrains are likely to preserve information on the isotopic composition of contemporary depleted mantle. □

Received 29 June; accepted 5 July 1994.

1. Moorbath, S., Welke, H. & Gale, N. H. *Earth planet. Sci. Lett.* **6**, 245-256 (1969).
2. Pidgeon, R. T. & Bowes, D. R. *Geol. Mag.* **109**, 247-258 (1972).
3. Chapman, H. J. & Moorbath, S. *Nature* **268**, 41-42 (1977).
4. Whitehouse, M. J. *Nature* **331**, 705-707 (1988).
5. Whitehouse, M. J. *Geochim. cosmochim. Acta* **53**, 717-724 (1989).
6. Corfu, F., Heaman, L. M. & Rogers, G. *Contr. Miner. Petrol.* (in the press).
7. Arth, J. G. & Hanson, G. N. *Contr. Miner. Petrol.* **37**, 161-174 (1972).
8. Arth, J. G. & Barker, F. *Geology* **4**, 534-536 (1976).
9. Rollinson, H. R. & Windley, B. F. *Contr. Miner. Petrol.* **72**, 265-281 (1980).
10. Weaver, B. L. & Tarney, J. *Earth planet. Sci. Lett.* **51**, 279-296 (1980).
11. Martin, H. *Geology* **14**, 753-756 (1986).
12. Whitehouse, M. J. *Tectonophysics* **161**, 245-256 (1989).
13. Cohen, A. S., O'Nions, R. K. & O'Hara, M. J. *Contr. Miner. Petrol.* **106**, 142-153 (1991).
14. Rollinson, H. R. & Fowler, M. B. in *Evolution of the Lewisian and Comparable Precambrian High Grade Terrains* (eds Park, R. G. & Tarney, J.) 57-71 (Spec. Publ. No. 27, Geol. Soc. Lond. 1987).
15. Rollinson, H. R. *Mineralog. Mag.* **51**, 345-355 (1987).
16. Burton, K. W. & O'Nions, R. K. *Earth planet. Sci. Lett.* **107**, 649-671 (1991).
17. Hamilton, P. J., Evensen, N. M., O'Nions, R. K. & Tarney, J. *Nature* **277**, 25-28 (1979).
18. Taylor, S. R. & McLennan, S. M. *The Continental Crust: Its Composition and Evolution* (Blackwell, Oxford, 1985).
19. Goldstein, S. L., O'Nions, R. K. & Hamilton, P. J. *Earth planet. Sci. Lett.* **70**, 221-236 (1984).
20. Waters, F. G., Cohen, A. S., O'Nions, R. K. & O'Hara, M. J. *Earth planet. Sci. Lett.* **97**, 241-255 (1990).
21. Chapman, H. J. *Nature* **277**, 642-643 (1979).
22. Heaman, L. M. & Tarney, J. *Nature* **340**, 705-708 (1989).
23. White, W. M. *Earth planet. Sci. Lett.* **115**, 211-226 (1993).
24. Asmerom, V. & Jacobsen, S. B. *Earth planet. Sci. Lett.* **115**, 245-256 (1993).
25. Heier, K. S. *Phil. Trans. R. Soc. A273*, 429-442 (1973).
26. Rudnick, R. L. & Presper, T. in *Granulites and Crustal Evolution* (eds Vielzeuf, D. & Vidal, P.) 523-550 (Kluwer, Amsterdam, 1990).
27. Tarney, J. & Weaver, B. L. in *Evolution of the Lewisian and Comparable Precambrian High Grade Terrains* (eds Park, R. G. & Tarney, J.) 45-56 (Spec. Publ. No. 27, Geol. Soc. Lond. 1987).
28. Göpel, C., Manhès, G. & Allègre, C. J. *Earth planet. Sci. Lett.* **121**, 153-171 (1994).
29. Steiger, R. H. & Jäger, E. *Earth planet. Sci. Lett.* **36**, 359-362 (1977).
30. Tatsumoto, M., Knight, R. J. & Allègre, C. J. *Science* **180**, 1279-1283 (1973).
31. Jacobsen, S. B. & Wasserburg, G. J. *Earth planet. Sci. Lett.* **50**, 139-155 (1980).
32. Fletcher, I. R. & Rosman, K. J. R. *Geochim. cosmochim. Acta* **46**, 1983-1987 (1982).

ACKNOWLEDGEMENTS. We thank H. J. Chapman, M. J. Bickle and M. J. Whitehouse for their thoughtful and constructive comments on this work, and S. J. G. Galer for insight into the U/Pb systematics. This research was supported by the UK NERC and the Royal Society.

Processing of tumour necrosis factor- α precursor by metalloproteinases

A. J. H. Gearing, P. Beckett, M. Christodoulou, M. Churchill, J. Clements, A. H. Davidson, A. H. Drummond, W. A. Galloway, R. Gilbert, J. L. Gordon*, T. M. Leber, M. Mangan, K. Miller, P. Nayee, K. Owen, S. Patel, W. Thomas, G. Wells, L. M. Wood & K. Woolley

British Biotech, Watlington Road, Cowley, Oxford OX4 5LY, UK

TUMOUR necrosis factor- α (TNF- α) is a potent pro-inflammatory and immunomodulatory cytokine implicated in inflammatory conditions such as rheumatoid arthritis, Crohn's disease, multiple sclerosis and the cachexia associated with cancer or human immunodeficiency virus infection¹. TNF- α is initially expressed as a 233-amino-acid membrane-anchored precursor which is proteolytically processed to yield the mature, 157-amino-acid cytokine². The processing enzyme(s) which cleave TNF- α are unknown. Here we show that the release of mature TNF- α from leukocytes cultured *in vitro* is specifically prevented by synthetic hydroxamic acid-based metalloproteinase inhibitors, which also prevent the release of TNF- α into the circulation of endotoxin challenged rats. A recombinant, truncated TNF- α precursor is cleaved to biologically active, mature TNF- α by several matrix metalloproteinase enzymes. These results indicate that processing of the TNF- α precursor is dependent on at least one matrix metalloproteinase-like enzyme, inhibition of which represents a novel therapeutic mechanism for interfering with TNF- α production.

Examination of the sequences surrounding the proposed cleavage site of the TNF- α precursor revealed homologies with peptide sequences known to be cleaved by matrix metalloproteinases (MMPs). We therefore tested the effects of several synthetic hydroxamic acid-based inhibitors of MMPs on the release of TNF- α . A number of these inhibitors prevented TNF- α production, whereas diastereoisomers, inactive against MMPs, did not. In an assay of TNF- α release from endotoxin-stimulated human blood cultures, BB-2284, BB-2275 and BB-2116 had IC₅₀ values (50% inhibitory concentrations) of 400, 580 and 230 nM respectively. This is approximately 100-fold more potent than pentoxifylline (see Fig. 1), which inhibits TNF- α production at the genetic level and is used in the clinic to treat diseases in which TNF- α is known to be involved³. Inhibitors of non-matrix metalloproteinases such as angiotensin converting enzyme, enkephalinase and endothelin converting enzyme^{4,5} had no effect on TNF- α production (Fig. 1), and inhibitors of serine proteases (aprotinin, 10 μ M; elastinal, 200 μ M), serine/cysteine proteases (leupeptin, 100 μ M), cysteine proteases (E64, 100 μ M), aspartyl proteases (pepstatin, 100 μ M) and aminopeptidase (Bestatin, 100 μ M)⁶ were also ineffective. The effect of BB-2284 was specific for TNF- α processing: no significant inhibition of the release of the interleukins IL-1 β , IL-6, IL-8 and IL-10, G-CSF or GM-CSF from whole blood cultures was observed (results for TNF- α , IL-1 β and IL-8 are shown in Fig. 2a-c).

To establish whether the processing of pro-TNF- α by a metalloproteinase was direct or indirect we used a recombinant glutathione-S-transferase-linked pro-TNF- α fusion protein (GST-TNF- α) as a substrate. This was recognized by several antibodies to TNF- α and was biologically active in an L-929 cytotoxicity assay⁷. Addition of MMPs such as PUMP (matrilysin; MMP-7), stromelysin-1 (MMP-3), collagenase (MMP-1),

* To whom correspondence should be addressed, at: NEURES Ltd, 4-10 The Quadrant, Barton Lane, Abingdon, Oxon OX14 3YS, UK.

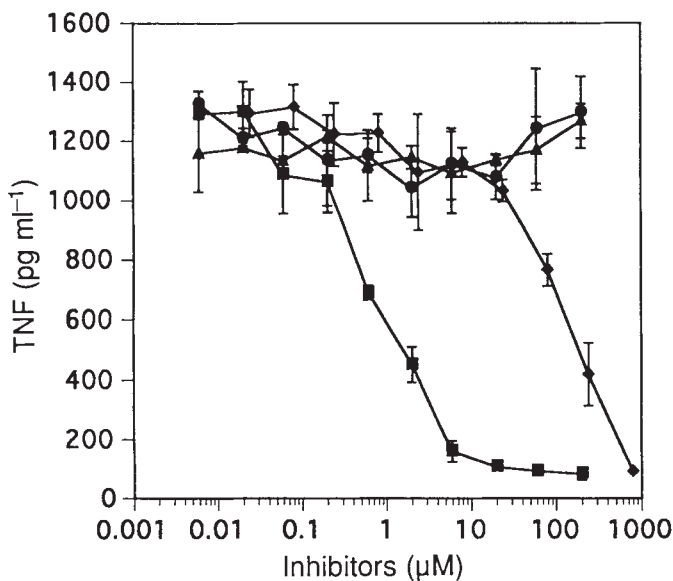


FIG. 1 Specific inhibitors of matrix metalloproteinases inhibit the release of mature TNF- α from human whole blood cultures. Control, unstimulated cultures produced no detectable TNF- α , whereas endotoxin stimulated the production of over $1,000 \text{ pg ml}^{-1}$. BB-2284 [3R-(3-tert-butoxycarbonyl-1S-methylcarbamoyl-propylcarbamoyl)-2S-(4-hydroxy-phenylsulphonylmethyl)-5-methyl-hexanohydroxamic acid] caused a dose-dependent inhibition of TNF- α release, with an IC_{50} value of 400 nM. Values of 230 and 580 nM respectively were obtained with the related compounds BB-2116 [3R-(5-acetylaminol-1S-methylcarbamoylpropylcarbamoyl)-2S-(4-hydroxy-phenylsulphonylmethyl)-5-methyl-hexanohydroxamic acid] and BB-2275 [2S-(4-hydroxy-phenylsulphonylmethyl)-3R-(3-methoxycarbonyl-1S-methylcarbamoyl-propylcarbamoyl)-5-methylhexanohydroxamic acid]. All structures are given in IUPAC nomenclature. Pentoxifylline (an inhibitor which acts on TNF- α mRNA¹⁶) was approximately 100-fold less potent than BB-2284. Captopril and phosphoramidon, which inhibit angiotensin converting enzyme and enkephalinase respectively, but not MMPs, had no effect on TNF- α release.

METHODS. Human peripheral blood was anticoagulated with heparin and diluted 1:4 with RPMI 1640 medium (Gibco). 1 ml cultures (in duplicate 24-well plates) were incubated for 18 hours at 37°C in the presence of 100 ng ml^{-1} LPS (Sigma) to stimulate TNF- α release¹⁷, with or without inhibitors. TNF- α concentrations in the culture supernatants were measured using a specific ELISA kit (R&D Systems).

92K or 72K gelatinases (MMP-9 and MMP-2) resulted in cleavage of GST-TNF- α ; the gelatinases were less potent than the other MMPs. A 17K M_r protein was produced which co-migrated with, and had identical specific activity to, an authentic TNF- α standard (Fig. 3). The N-terminal sequence of the 17K cleavage product was identical to mature TNF. This cleavage was abolished by BB-2284. Although the fusion protein was cleaved by thermolysin (a bacterial metalloprotease) and by trypsin, the cleavage products generated by these enzymes migrated at lower M_r than the authentic TNF- α standard and were 10-fold less active in the L929 TNF- α bioassay; this cleavage was not blocked by BB-2284. These experiments confirmed that biologically relevant processing of pro-TNF- α is achieved by MMPs and that, *in vitro*, production of mature, secreted TNF- α can be blocked by specific inhibitors of these enzymes.

Our inhibitors of MMPs also prevented TNF- α release *in vivo*. An intravenous bolus of endotoxin induces a characteristic pattern of TNF- α release in rats: TNF- α blood levels rise rapidly, peaking at 60 min, and returning to baseline over the next 2 hours. An intravenous infusion of BB-2275, commencing 15 min

before the endotoxin administration, inhibited the appearance of TNF- α in a dose-dependent manner (Fig. 4a). The release of TNF- α into the blood was also inhibited by a bolus injection (Fig. 4b), even when this was given 50 min after endotoxin injection, when the rate of accumulation of TNF- α is at its peak (Fig. 4c), which suggests that the inhibitors are acting at a terminal step in the TNF- α production pathway.

MMPs thus appear to be critical components of TNF- α release both *in vitro* and *in vivo*. Although an undiscovered TNF- α converting metalloenzyme may exist in some cell types, pro-TNF- α processing can be accounted for by co-localization of known MMPs. If this processing takes place on the cell surface, blocking it could theoretically lead to accumulation of membrane-anchored pro-TNF- α , which is biologically active. However, we saw no such accumulation—probably because there is a homeostatic feedback mechanism which ensures that pro-TNF- α is rapidly recycled inside cells⁸. MMP action is controlled by natural tissue inhibitors of metalloproteinases (TIMPs)^{9,10}, often co-produced with MMPs, which prompted us to test the effect of TIMP-2 on TNF- α release. At concentrations which

FIG. 2 Cytokine specificity of metalloproteinase inhibitors. Whole blood cultures were set up as described in Fig. 1. Cytokine levels were measured in the culture supernatants at intervals over 48 h using specific ELISA kits (R&D Systems). a, TNF- α was detectable within 4 h of stimulation with endotoxin, peaked at 24 h and had returned to basal levels by 48 h (■). In the presence of $100 \mu\text{M}$ BB-2284 (●), TNF- α remained at basal levels for the duration of the culture. b, IL1 β increased with time, reaching a plateau at 24 h which was sustained up to 48 h (■). BB-2284 ($100 \mu\text{M}$) had no significant effect on IL1 β levels (●). c, IL-8 levels rose continuously from 8 to 48 h (■) and were not reduced by $100 \mu\text{M}$ BB-2284 (●).

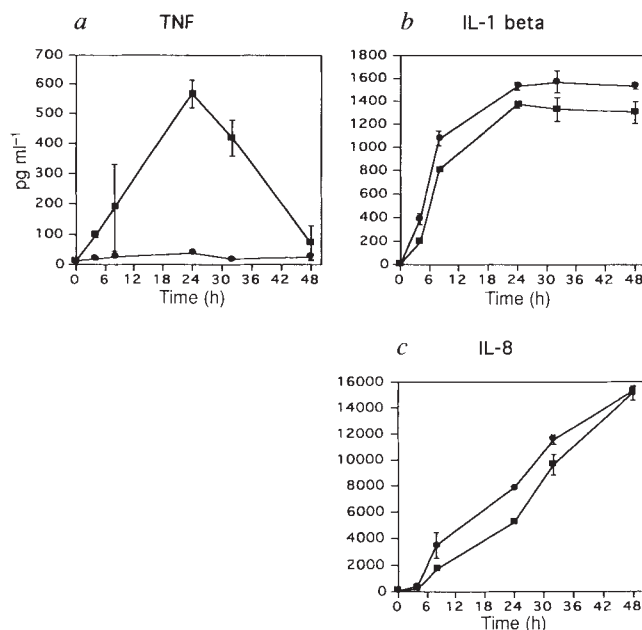
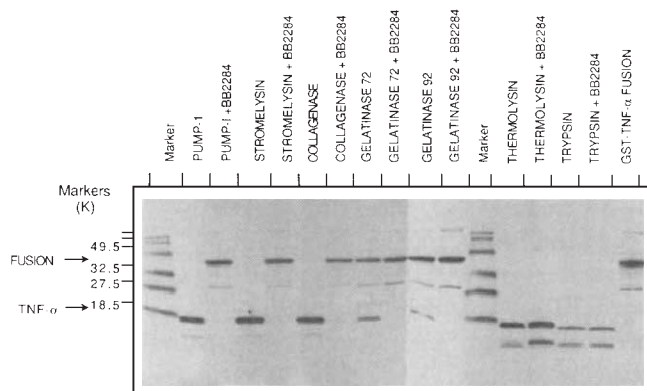


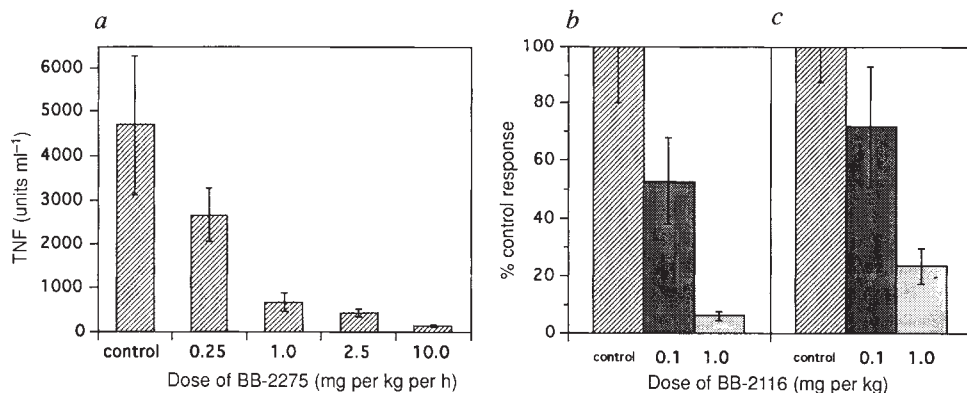
FIG. 3 Recombinant pro-TNF- α substrate is cleaved to mature TNF- α by purified MMPs. Western blotting with anti-TNF- α antibody was used to follow the cleavage. The substrate was cleaved by PUMP, stromelysin-1, collagenase, 72K- and 92K-gelatinase to produce 17K TNF- α , with the first three enzymes particularly effective. The cleavage by MMPs was completely blocked by BB-2284. The bacterial metalloprotease, thermolysin, cleaved the substrate to 17K and 14K bands, but neither band was authentic TNF- α and this cleavage was not prevented by BB-2284. Similarly, cleavage by trypsin, which gave bands at 16K and 13.5K, was unaffected by BB-2284. The cleavage products from thermolysin and trypsin were 10-fold less potent in the L929 assay than products from the MMPs.

METHODS. The recombinant pro-TNF- α substrate was made from a truncated, precursor TNF- α lacking the cytoplasmic and membrane-spanning sequences, constructed by cloning a sequence encoding HFGVIGPQREESPRDLISLPLAQA onto the 5' end of the mature TNF- α gene (R&D Systems). This precursor gene was transferred into the pGEX-2T (Pharmacia) expression vector in the same translational reading frame as the glutathione S-transferase (GST) gene. The fusion protein was expressed in *E. coli* strain HW87 following induction with 0.1 M isopropyl- β -thiogalactopyranoside. The cells were disrupted by repeated freeze-thawing, and the fusion protein was purified on glutathione-agarose beads according to the manufacturer's instructions (Pharmacia). The eluted GST-TNF- α fusion protein was detected in a specific ELISA for TNF- α , and was biologically active in an L929 cytotoxicity assay⁷. To examine the effects of enzymes on the processing of the TNF- α substrate an *in vitro* cleavage assay was devised. 5 μ g of the GST-TNF-



α bound to glutathione beads was incubated at 37 °C for 2 h with purified, activated matrix metalloproteinases, or other enzymes, as follows: 0.04 units recombinant human PUMP¹⁸; 0.02 units recombinant human stromelysin-1¹⁸; 0.2 units human fibroblast collagenase¹⁹; 13 units 72K-gelatinase²⁰; 1 unit human melanoma-derived 92K-gelatinase²⁰; 0.1 unit of thermolysin; 2 μ g of trypsin. The reaction mixtures were analysed by SDS-PAGE, followed by western blotting with a mouse monoclonal antibody to TNF- α (Endogen Inc.).

FIG. 4 Synthetic inhibitors of MMPs prevent the release of TNF- α into the blood of endotoxin-treated rats. TNF- α levels, which are negligible before endotoxin, rise to a peak after 60 min; all TNF- α levels were measured at this time. An intravenous infusion of BB-2275, commencing 15 min before endotoxin, inhibited the production of TNF- α in a dose-dependent manner (a). An intravenous bolus of BB-2116, given 30 min (b) or 50 min (c) after endotoxin, inhibited TNF- α production at doses of 0.1 mg per kg body weight and above; peak TNF- α levels in control animals were 2,800 and 3,700 pg ml⁻¹ respectively in these experiments.



METHODS. Male Sprague-Dawley rats weighing 400–500 g were anaesthetized with sodium pentobarbitone/thiopental given intraperitoneally. Endotoxin (0.5 mg per kg) was given as a bolus injection via the right jugular vein, and BB-2275 or BB-2116 via the left jugular vein. Blood samples were taken from the carotid artery, and TNF- α levels were measured in serum using an ELISA

(Genzyme) designed to measure murine TNF- α , which cross reacts with rat TNF- α . Results are expressed as units per ml compared to the murine TNF- α standard, where 1 unit is the equivalent of 1 pg murine TNF- α .

abolished the activity of gelatinase or collagenase, TIMP-2 had no effect, though much higher concentrations ($>1 \mu$ M) were partially inhibitory. The TIMPs are much less potent against truncated MMPs¹¹ such as matrilysin, which suggests that a similar enzyme could play a role in pro-TNF- α processing. In conditions such as arthritis, where matrix degradation is a key feature, TIMP and TNF- α antibodies are both active^{12–14}. Syn-

thetic inhibitors like those used in our studies are very effective in animal models¹⁵, suggesting that drugs which block both TNF- α and MMPs should also be effective in other diseases where TNF- α production is an important feature of the pathology.

While this paper was under review, a report of another metalloproteinase inhibitor that blocks TNF- α release and protects mice from a lethal dose of endotoxin has been published²¹. □

Received 9 June; accepted 20 July 1994

- Vassalli, P. A. *Rev. Immun.* **10**, 411–452 (1992).
- Kriegler, M., Perez, C., Defay, K., Albert, I. & Lu, S. D. *Cell* **53**, 45–53 (1988).
- Dezube, B. J. et al. *Cancer Immun. Immunother.* **36**, 57–60 (1993).
- Fulcher, I. S. et al. *Biochem. J.* **203**, 519–522 (1982).
- Ikegawa, R. et al. *Biochem. biophys. Res. Commun.* **171**, 669–675 (1990).
- Gee, N. S. & Kenny, A. J. *Biochem. J.* **246**, 97–102 (1987).
- Matthews, N. & Neale, M. L. in *Lymphokines and Interferons: A Practical Approach* (eds Clemens, M. J., Morris, A. G. & Gearing, A. J. H.) 221–225 (IRL, Oxford, 1987).
- Pradines-Figueras, A. & Raetz, C. R. A. *J. Biol. Chem.* **267**, 23261–23268 (1992).
- Docherty, A. J. P. et al. *Nature* **318**, 66–69 (1985).
- Stetler-Stevenson, W. G. et al. *J. Biol. Chem.* **264**, 17374–17378 (1989).
- Fridman, R. et al. *J. Biol. Chem.* **267**, 15398–15405 (1992).

- Carmichael, D. F. et al. *Agents Actions* **27**, 378–379 (1989).
- Williams, R. O. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 9784–9788 (1992).
- Elliot, M. J. et al. *Arthritis Rheumat.* **36**, 1681–1690 (1993).
- Dimartino, M. J. et al. *J. cell. Biochem.* **51**, 179 (1991).
- Endres, S. et al. *Immunology* **72**, 56–60 (1991).
- Wilson, B. M. G. et al. *J. Immun. Meth.* **139**, 233–240 (1991).
- Cawston, T. E. et al. *Biochem. J.* **195**, 159–165 (1981).
- Cawston, T. E. & Barrett, A. J. *Analyt. Biochem.* **99**, 340–345 (1979).
- Seliars, A. et al. *Biochem. J.* **171**, 493–496 (1978).
- Mohler, K. M. et al. *Nature* **370**, 218–220 (1994).

ACKNOWLEDGEMENTS. We thank all members of the British Biotech metalloproteinase programme for their contribution to this work. We also thank G. Smith and C. Greasley for assisting with preparation of the manuscript.

Regulation of tumour necrosis factor- α processing by a metalloproteinase inhibitor

Gerard M. McGeehan*, J. David Becherer*, Robert C. Bast Jr†, Cinda M. Boyer†, Brian Champion*, Kevin M. Connolly*, James G. Conway*, Paul Furdon*, Stephen Karp‡, Sudha Kidao§, Andrew B. McElroy*, James Nichols*, Katherine M. Pryzwansky§, Frank Schoenen*, Les Sekut*, Anne Truesdale*, Margrith Verghese*, Janet Warner* & Judy P. Ways*

* Glaxo Inc. Research Institute, Research Triangle Park, North Carolina 27709, USA

† Department of Medicine, Duke Comprehensive Cancer Center, Duke University Medical Center, Durham, North Carolina 27710, USA

‡ Department of Oncology, Jewish General Hospital, 3755 Chemin de Cote-Ste-Catherine, Montreal, Quebec H3T 1E2, Canada

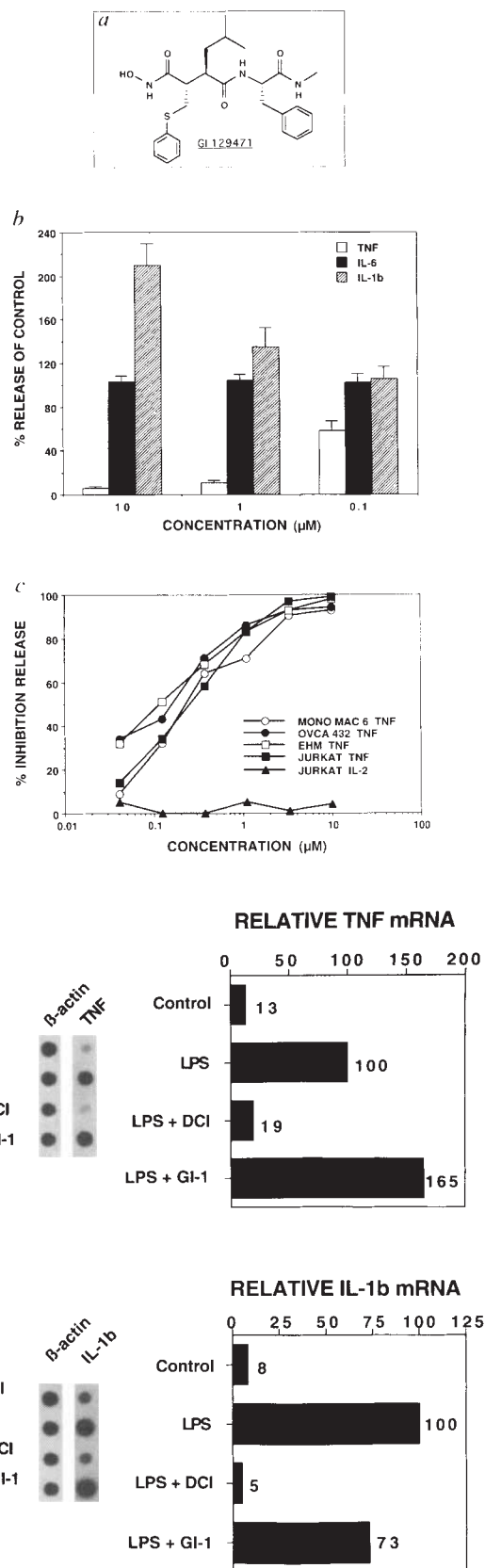
§ Department of Pathology, University of North Carolina School of Medicine, Chapel Hill, North Carolina 27516, USA

TUMOUR necrosis factor- α (TNF- α) is a potent pro-inflammatory agent produced primarily by activated monocytes and macrophages¹. TNF- α is synthesized as a precursor protein of *M*_r 26,000 (26K) which is processed to a secreted 17K mature form by cleavage of an Ala-Val bond between residues 76–77. The enzyme(s) responsible for processing pro-TNF- α has yet to be identified. Here, we describe the capacity of a metalloproteinase inhibitor, GI 129471, to block TNF- α secretion both *in vitro* and *in vivo*. The inhibition is specific to TNF- α ; the production of other secreted cytokines, such as the interleukins IL-1 β , IL-2, or IL-6, is not inhibited. The mechanism of inhibition occurs at a post-translational step in TNF- α production. Our data suggest that TNF- α processing is mediated by a unique Zn²⁺ endopeptidase which is inhibited by GI 129471 and would represent a novel target for therapeutic intervention in TNF- α associated pathologies.

The proteinase inhibitor GI 129471 (Fig. 1a) was tested for inhibition of cytokine production from lipopolysaccharide

FIG. 1 GI 129471 specifically inhibits TNF- α secretion at a post-transcriptional step. a, Structure of GI 129471. b, Specificity of cytokine inhibition. Compound was tested at three concentrations (10, 1.0, and 0.1 μ M) and was added 15 min prior to stimulation with LPS (10 ng ml⁻¹). Results are presented as per cent of the control (no compound). c, Dose responsive inhibition of TNF- α release from activated human cell lines by GI 129471. Levels of secreted TNF- α were determined by ELISA. IC₅₀ values were calculated using standard curve fitting programs. d, The effect of GI 129471 and DCI on IL-1 β and TNF- α mRNA accumulation in LPS-treated THP-1 cells. Values were plotted as a relative per cent of the LPS sample for TNF- α and IL-1 β . GI-1 refers to GI 129471.

METHODS. b, c, Cells were maintained at 37 °C and 5% CO₂ in RPMI 1640 media supplemented with 2 mM L-glutamine, 50 units ml⁻¹ penicillin, 50 μ g ml⁻¹ streptomycin in either 10% fetal calf serum (cell lines) or 2.5% heat-inactivated pooled human serum (PBMC). Lymphocytic and monocytic lines: PBMC were stimulated with LPS (10 ng ml⁻¹). Mono Mac 6 cells were stimulated by addition of LPS (10 ng ml⁻¹) and PMA (30 ng ml⁻¹). EHM and Jurkat were stimulated with ionomycin (2 μ M) and PMA (20 ng ml⁻¹). Culture media were removed at 4 h (Mono Mac 6) or 16 h (EHM, Jurkat, PBMC) and cytokine levels were determined by ELISA. Viability was >95 % (LDH assay kit; Sigma, St. Louis, MO). OVCA 432: Cells were plated in flat bottom microtitre plates at 2.5 $\times$ 10⁵ ml⁻¹ and incubated overnight. The medium was removed and fresh media containing different concentrations of GI 129471 were added. Cells were stimulated with IL-1 β (20 ng ml⁻¹), media were



removed after 24 h and TNF- α concentrations in the supernatants were determined by ELISA. d, Total RNA was isolated⁵ from THP-1 cells following treatment with LPS (500 ng ml⁻¹), LPS + dichloroisocumarin (DCI) (100 μ M) or LPS + GI 129471 (10 μ M). RNA was analysed by dot blot using either ³²P-labelled TNF- α or IL-1 β specific probes. The c.p.m. were normalized to c.p.m. values obtained after hybridization to a β -actin probe.

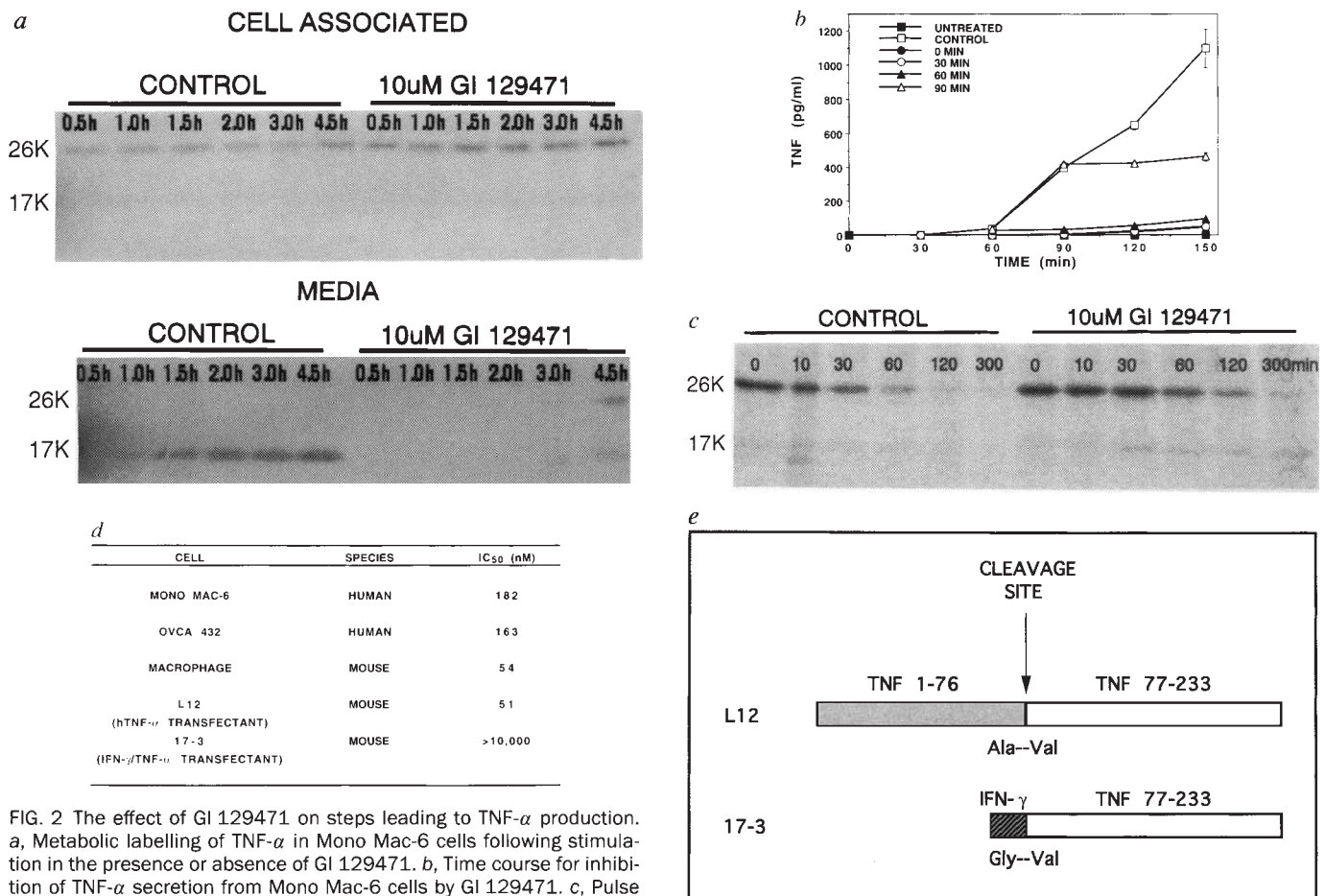


FIG. 2 The effect of GI 129471 on steps leading to TNF- α production. **a**, Metabolic labelling of TNF- α in Mono Mac-6 cells following stimulation in the presence or absence of GI 129471. **b**, Time course for inhibition of TNF- α secretion from Mono Mac-6 cells by GI 129471. **c**, Pulse chase experiment in the presence or absence of GI 129471. **d**, Comparison of the IC₅₀ values of GI 129471 on TNF- α secretion from human or mouse cells. **e**, Schematic representation of the human TNF- α constructs used to transduce the murine fibrosarcoma, 205 F4¹⁶. **METHOD.** **a, c**, For metabolic labelling, Mono Mac-6 cells were activated (LPS/PMA) in the presence of ³⁵S-cysteine (75 μ Ci ml⁻¹) containing media with or without GI 129471 (10 μ M). For pulse-chase, cells were stimulated (LPS/PMA) and metabolically labelled with ³⁵S-cysteine (75 μ Ci ml⁻¹) for 1 h. Following the pulse, cells were washed and resuspended in complete media with or without GI 129471. At specific intervals aliquots were removed, cells and media were separated, and TNF- α was immunoprecipitated as described¹¹. Proteins were separated on gradient gels. Gels were scanned by Phosphorimager (Molecular

(LPS) stimulated human peripheral blood monocytes (PBMC). This compound gave a dose responsive inhibition of TNF- α release with an estimated IC₅₀ of 180 nM (Fig. 1b). In contrast, IL-6 secretion was not affected, while a twofold stimulation of IL-1 β release occurred at the highest dose (10 μ M). The mechanism of IL-1 β stimulation is unclear, but may reflect an effect of TNF- α on secretion of IL-1 β by monocytes. The same inhibition profile for cytokine production also was observed for the monocytic lines Mono Mac-6 (ref. 2) and THP-1 (data not shown). GI 129471 was tested in other human cell types to determine if the effect was restricted to monocytes. TNF- α secretion was induced in Mono Mac-6 cells, a Jurkat T-cell line³, the EHM B-cell line and the ovarian tumour line, OVCA 432 (ref. 4), using various stimuli. TNF- α secretion was inhibited similarly in all cell types (IC₅₀ ~120–220 nM, Fig. 1c). In contrast, release of IL-1 β , IL-2 or IL-6 was not inhibited (IL-2 shown in Fig. 1c). The secretion of several other membrane anchored proteins that require proteolytic processing (CSF-1, TGF- α and transferin receptor) also was not significantly affected (data not shown). Therefore, GI 129471 specifically inhibits TNF- α release in all cell types tested regardless of the stimulus used to initiate production.

Dynamics). **b**, Mono Mac-6 cells were plated at 1×10^6 ml⁻¹ and stimulated with LPS/PMA. Compound was added at $t = 0, 30, 60$ and 90 min. Media were removed at 150 min and TNF- α levels were determined by ELISA. **d**, Sodium periodate elicited mouse peritoneal macrophages were prepared as described¹⁷. Cells were plated at 5×10^5 ml⁻¹ and stimulated with LPS (2.5 ng ml⁻¹). After 20 h murine TNF- α and IL-6 were quantitated by ELISA analysis. The cell lines L12 and 17-3 were maintained in RPMI as described¹⁶. For experiments with GI 129471, cells were trypsinized, resuspended in fresh media either with or without compound at 5×10^5 cells ml⁻¹ and plated in 24 well plates. Media were removed at 24 h and human TNF- α levels were determined by ELISA.

GI 129471 did not inhibit TNF- α or IL-1 β message accumulation in LPS stimulated THP-1 cells (Fig. 1d). This indicates GI 129471 does not inhibit LPS signalling or TNF- α gene transcription. In contrast, dichloroisocoumarin (DCI) abolished both TNF- α and IL-1 β message production. DCI and other serine protease inhibitors previously have been reported to block TNF- α production^{6,9}. These inhibitors are now known to block activation of the transcription factor NF- κ B and subsequent cytokine gene activation¹⁰. GI 129471 acts at a post-transcriptional step and, therefore, inhibits TNF- α production by a distinct mechanism. The specificity for inhibition of TNF- α release suggests that a common but unique target is affected in TNF- α producing cells.

The effect of GI 129471 on TNF- α protein synthesis is shown in Fig. 2a. Mono Mac-6 cells were metabolically labelled during stimulation and TNF- α was immunoprecipitated from the media or cell lysates. In the absence of GI 129471, 17 K mature TNF- α appeared in the media at 1 h and increased over time. The 26 K pro-TNF- α was found in the cell associated fraction and reached steady state levels at 30–60 min. Cells labelled in the presence of GI 129471 secreted almost no TNF- α while there was a twofold increase in the steady state level of cell associated TNF- α , which

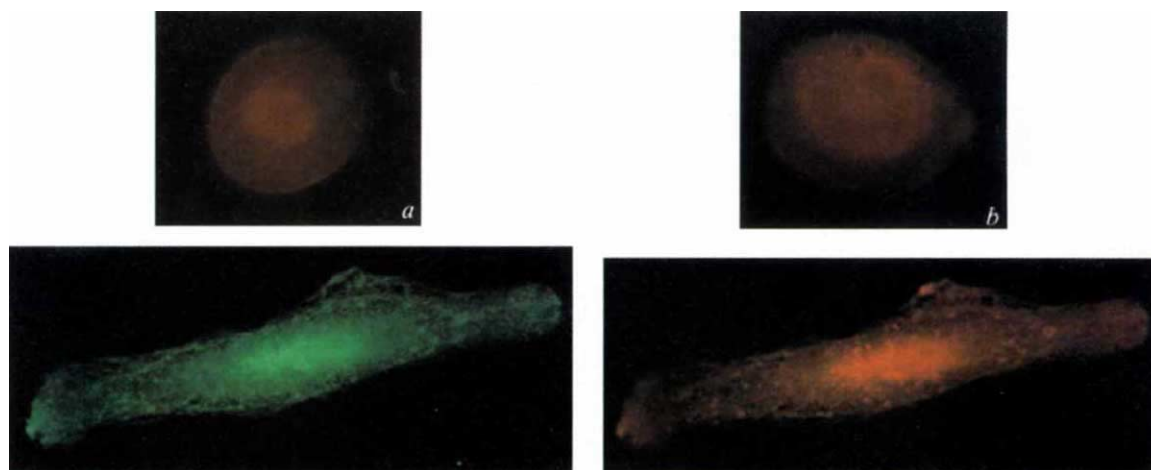


FIG. 3 Immunolocalization of TNF- α . Upper panel, Human monocyte derived macrophages were incubated in the presence of GI 129471 alone (a) or LPS alone (b). Cells were fixed and stained for TNF- α . Lower Panel. Cells were stimulated with LPS (10 ng ml^{-1}) in the presence of GI 129471 ($10 \mu\text{M}$). Cells were fixed and double immunofluorescence microscopy was performed using anti β -COP (a) and anti TNF- α (b) antibodies.

METHODS. Mononuclear leukocytes were isolated from human peripheral blood and were purified by elutriation (>97% pure). Cells were resuspended at $1 \times 10^6 \text{ cells ml}^{-1}$ in M199 media containing 0.1% autologous serum and were maintained on coverslips at 37°C in M199

media in the presence of 2.5% autologous serum in 5% CO_2 . After 6 days, cells had differentiated into macrophages as determined by increased release of β -galactosidase. For localization studies, cells were treated with GI 129471 ($10 \mu\text{M}$) or LPS (10 ng ml^{-1}) for 4 h. Cells were fixed and incubated at 4°C overnight with mouse anti-human TNF- α monoclonal antibody (Boehringer Mannheim), washed with PBS, stained for 30 min with TRITC goat anti-mouse IgG (Cappel, Durham, NC) and washed again. For dual fluorescence, cells were also incubated with rabbit anti- β -COP IgG (gift from J. Lippincott-Schwartz, NIH) and a FITC goat anti-rabbit (Fab') $_2$ IgG (Cappel, Durham, NC).

was predominantly in the 26K proform. This result was confirmed by an enzyme-linked immunosorbent assay (ELISA) specific for TNF- α (data not shown). Furthermore, a rapid onset of action of GI 129471 was demonstrated in a time-course experiment (Fig. 2b). TNF- α secretion was inhibited when compound was added to the media before or after protein production had begun. When added 90 min after stimulation, export of additional TNF- α was still blocked greater than 90%. Together these data are consistent with the inhibitor acting at a post-translational step in TNF- α production.

Cell surface staining of stimulated Mono Mac-6 cells showed a small increase in surface TNF- α (4–7-fold) in the presence of GI 129471 (data not shown). However, this small increase cannot account for the overall twofold increase in cell associated TNF- α ¹². The effect of GI 129471 on cell-associated TNF- α was evaluated in a pulse-chase experiment. In the absence of inhibitor, levels of 26K pro-TNF- α decreased with a $t_{1/2}$ of 15–20 min (Fig. 2c). Similarly, the processed 17K TNF- α was secreted into the media with a similar half-life (data not shown). These data agree with the reported half-life for pro-TNF- α secretion from activated Mono Mac-6 cells¹¹. In contrast, in the presence of GI 129471 intracellular levels of 26K pro-TNF- α remained steady for 30 min and then decreased with $t_{1/2}$ of 20–30 min. This decrease is likely due to proteolysis as little TNF- α was secreted into the medium (data not shown). These results indicate that GI 129471 acts at a step distal to protein synthesis and may affect processing of TNF- α by inhibition of a specific convertase.

GI 129471 was originally designed as a $\text{P}_1\text{-P}_2'$ Leu-Phe analogue inhibitor of collagenase¹³. As human TNF- α contains valine at P_1 ¹⁴ we reasoned that if GI 129471 binds to a specific TNF- α convertase, it would be a more potent inhibitor in cells where the endogenous TNF- α substrate contains a P_1 leucine. Mouse pro-TNF- α contains a P_1 leucine residue at the cleavage site¹⁵. As shown in Fig. 2d, GI 129471 is a 3–4-fold more potent inhibitor in mouse macrophages ($\text{IC}_{50} \sim 54 \text{ nM}$) than human cells ($\text{IC}_{50} \sim 180 \text{ nM}$; Mono Mac-6). Mouse cell lines L12 and 17-3 were utilized to demonstrate further that GI 129471 inhibits cellular events unique to TNF- α processing. These lines constitu-

tively express wild-type human TNF- α , L12, or a chimaeric form of TNF- α fused to the IFN- γ signal sequence, 17-3 (ref. 16). 17-3 contains a Gly-Val cleavage site adjacent to the IFN- γ signal sequence making it a substrate for signal peptidase, which does not process human TNF- α ¹⁸. GI 129471 inhibited release of TNF- α from the L12 cell line with an IC_{50} (51 nM) similar to that seen for murine macrophages (Fig. 2e). In contrast, GI 129471 did not inhibit TNF- α production by cell line 17-3. These results indicate that GI 129471 has no effect on the

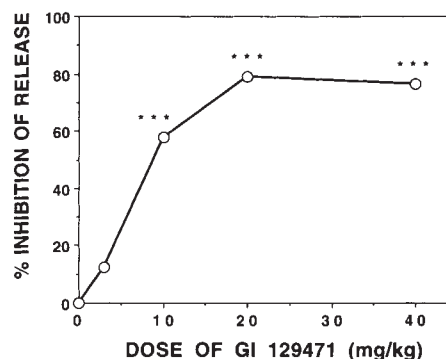


FIG. 4 Dose responsive inhibition of LPS-induced TNF- α production *in vivo* by compound GI 129471. Doses of GI 129471 were given subcutaneously 10 min before LPS injection. TNF- α was measured in the plasma 1.5 h after LPS administration. Data represent per cent inhibition compared to vehicle-treated controls. (***) Signifies $P < 0.001$ compared to the vehicle treated controls using a 2-sided *t*-test.

METHODS. Compound was dissolved in DMSO and added to a mixture of 0.9% NaCl and 30% Trappsol HPB-20 (CTD Inc., Gainesville, FL) for a final DMSO concentration of 1%. Compound was injected s.c. in a volume of 0.2 ml 10 min before LPS injection. C3H/He female mice were injected i.p. with 200 μg per kg body weight of LPS in PBS and were killed 90 min after LPS injection by CO_2 asphyxiation. Blood was taken immediately from the inferior vena cava, plasma was prepared and frozen at -80°C . Plasma concentrations of TNF- α were measured by ELISA.

trafficking of the mature form of TNF- α in mouse cells and that it affects a convertase distinct from signal peptidase. Further, the enhanced inhibition in mouse cells is consistent with GI 129471 acting directly on a novel Zn²⁺ endopeptidase responsible for processing of TNF- α .

Immunolocalization studies were conducted in human monocyte derived macrophages to determine the subcellular localization of TNF- α in the presence of GI 129471. Treatment with GI 129471 alone gave background staining for TNF- α while LPS stimulation produced diffuse staining throughout the cell (Fig. 3 upper panel). In contrast, in the presence of LPS and GI 129471, TNF- α was localized to perinuclear microsomes (Fig. 3 lower panel b). Colocalization with reagents specific to subcellular markers showed that staining with antisera specific for Golgi gave a pattern identical to the TNF- α antibody staining (Fig. 3 lower panel a). This indicates TNF- α was sequestered in the Golgi in the presence of GI 129471.

GI 129471 was shown to inhibit TNF- α production *in vivo* using a murine model of LPS induced TNF- α release. In two initial experiments, GI 129471 given subcutaneously at 40 mg per kg body weight reduced serum TNF- α levels by 88% and 87%, respectively (data not shown). Dose responsive inhibition of TNF- α production yielded an IC₅₀ between 3–10 mg per kg (Fig. 4).

The results presented here demonstrate that GI 129471 is a potent and selective inhibitor of TNF- α secretion both *in vitro* and *in vivo*. Given that GI 129471 was designed as a metalloproteinase inhibitor¹³, the data suggest TNF- α convertase is a Zn²⁺ dependent endopeptidase. This compound targets the active site of the MMPs which contain the highly conserved Zn²⁺ binding motif, HEXGH. This motif is found in many neutral proteinases^{19–22} including several recently described precursor processing enzymes^{23–25}. GI 129471 and related analogues possess anti-inflammatory activity in animal arthritis models (unpublished results). These effects *in vivo* were ascribed originally to inhibition of MMP activity; however, given the effect on TNF- α processing, the activity *in vivo* may be due to inhibition of TNF- α release.

While this paper was under review, Mohler *et al.*²⁶ described another metalloproteinase inhibitor that blocks TNF- α release and protects mice from a lethal dose of endotoxin. □

Regulation of cell number in *Drosophila*

Ana Busturia & Peter A. Lawrence

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

DURING animal development, different parts grow independently (such as the left and right hands) but they stop growing when they reach the correct size. In most insects, growth of the epidermis is so controlled that, at each moult, there is a precise and proportionate increase in cell number¹. The mechanisms responsible for this size regulation are not known², but rigid programming of the number of cell divisions is not a requirement as even sister cells in an epithelial sheet divide variably^{3–5}. In the abdomen of dipterans, such as *Drosophila*, the opportunity for regulation is limited, because mitoses occur only in the embryo and during metamorphosis and not during larval growth. Here we used embryos with a reduced number of cells in the abdominal primordia to determine whether they can regulate towards the normal during subsequent growth. In contrast to expectations^{6–8}, we find no evidence for regulation of cell number.

The ground plan of the *Drosophila* embryo is first mapped out by a morphogen gradient consisting of the *bicoid* protein⁷. The sets of cells that make the parasegments are allocated at about the time of gastrulation⁹, whereas the segment borders are probably fixed somewhat later¹⁰. Following gastrulation there are about three rounds of cell division in the embryo¹¹ and, after these, the larval epidermis grows only by an increase in the size of individual cells. Later, the histoblasts divide about nine times to give rise to the adult abdominal epidermis¹².

We looked for upregulation, by using a stock that carried seven copies of the *bicoid* gene (*7 bcd*); this is the maximum number of doses that can be used¹³. These embryos have an excess of the *bicoid* morphogen; they have an enlarged head and the parasegments of the trunk are crowded into a reduced space⁷. At the gastrulation stage, when the parasegments are demarcated by the stripes of expression of *even-skipped* and *fushi tarazu*¹⁴ the embryos are all alike (Fig. 1a). Subsequently, many develop and, although some show segmental abnormalities, about a quarter hatch to give larvae (Fig. 1 legend). The most severely affected parasegment is number 8 (corresponding to the third abdominal segment or A3 (ref. 2) and counts show a reduction in cell number of 31% in the gastrula (Fig. 1b). Much later, in the third-stage larva, the cell number in A3 is still 35% less than in the wild type, showing that there has been no significant upregulation in the embryo.

This failure to regulate also applies to the imaginal histoblasts of the abdomen¹⁵; for, in *7 bcd*, each of the three groups is reduced in number of cells by about 40% (Fig. 1f). Later, at metamorphosis, these histoblasts go through about nine rounds of cell division¹² and we expected some regulation to occur then. However, although we were unable to count every adult cell, there was a significant shortfall of 25% in the number of bristles (which gives a measure of the number of cells; Fig. 1 legend) in A3 of the *7 bcd* adults, indicating a complete or partial lack of regulation.

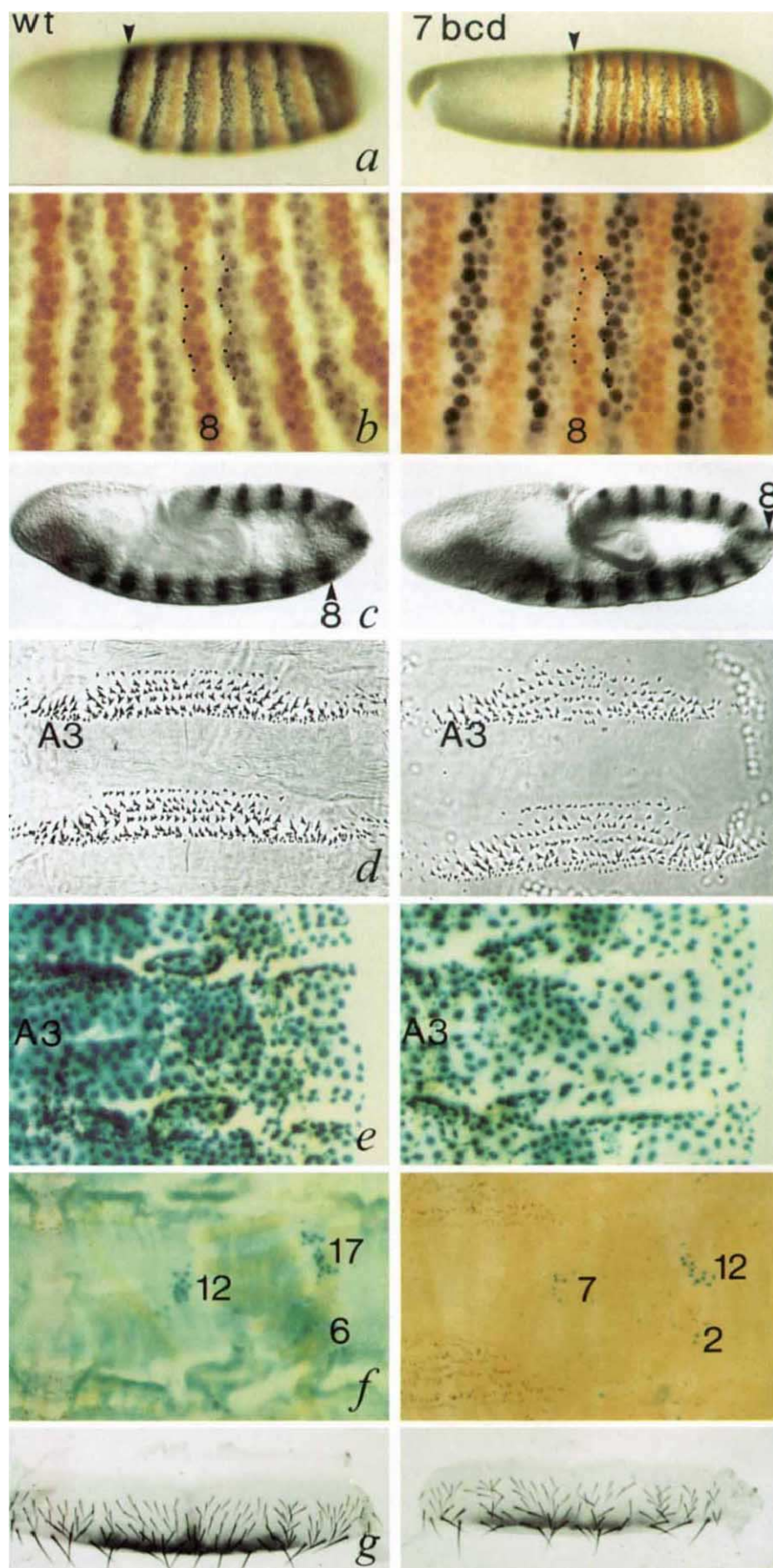
Can segments with reduced numbers of cells make normal patterns? It seems not, or at least not completely. In the larva, the pattern of denticles in the segment that is most severely affected (A3) is frequently disturbed and the denticle bands are often fused to those in neighbouring segments. Even when the bands are intact they are usually imperfect and often lack the most anterior row (Fig. 1d). In the adult there is a loss of bristles and the pigment band is reduced (Fig. 1g).

We have reduced the number of cells in the primordia by another method, that is by irradiating wild-type embryos during

Received 17 June; accepted 26 July 1994.

- Vassalli, P. A. *Rev. Immun.* **10**, 411–452 (1992).
- Ziegler-Heitbrock, H. W. L., Thiel, E., Futterer, A., Herzog, V., Wirtz, A. & Reithmüller, G. *Int. J. Cancer* **41**, 456–461 (1988).
- Bierer, B. *et al. Proc. natn. Acad. Sci. U.S.A.* **87**, 9231–9235 (1990).
- Bast, R. C. *et al. J. clin. Invest.* **68**, 1331–1337 (1981).
- Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294 (1979).
- Black, R. A., Greenstreet, T. A. & March, C. J. *Cytokine* **3**, 480 (1992).
- Scuderi, P. J. *Immun.* **143**, 168–173 (1989).
- Niehorster, M., Tiegs, G., Schade, U. F. & Wendel, A. *Biochem. Pharmac.* **40**, 1601–1603 (1990).
- Kim, K.-U., Kwon, O.-J. & Jue, D.-M. *Immunology* **80**, 134–139 (1993).
- Henkel, T. *Nature* **365**, 182–185 (1993).
- Pradines-Figueres, A., Raetz, C. R. H. *J. biol. Chem.* **267**, 23261–23268 (1992).
- Jue, D. M., Sherry, B., Luedke, C., Manogue, K. R. & Cerami, A. *Biochemistry* **29**, 8371–8377 (1990).
- Campion, C., Dickens, J. P., Crimmin, M. J. Patent WO 90/05719, British Biotechnology, 1990.
- Pennica, D. *et al. Nature* **312**, 724–729 (1984).
- Pennica, D., Hayflick, J. S., Bringman, T. S., Palladino, M. A. & Goeddel, D. V. *Proc. natn. Acad. Sci. U.S.A.* **82**, 6060–6064 (1985).
- Karp, S. E. *et al. J. Immun.* **149**, 2076–2081 (1992).
- Verghese, M. W. & Snyderman, R. *J. Immun.* **127**, 288–293 (1981).
- Müller, R., Marmenout, A. & Fiers, W. *FEBS Lett.* **197**, 99–104 (1986).
- Valle, B. L. & Auld, D. S. *Biochemistry* **29**, 5647–5659 (1990).
- Lovejoy, B. *et al. Science* **263**, 375–377 (1994).
- Bode, W., Gomis-Ruth, F. X., Huber, R., Zwilling, R. & Stocker, W. *Nature* **358**, 164–166 (1992).
- Matthews, B. W., Jansonius, J. N., Colman, P. M., Schoenborn, B. P. & Dupourque, D. *Nature new Biol.* **238**, 37 (1972).
- Pierotti, A., Dong, K.-W., Glucksmann, M. J., Orlowski, M. & Roberts, J. *Biochemistry* **29**, 10323–10329 (1990).
- Isaya, G., Kalousek, F. & Rosenberg, L. E. *Proc. natn. Acad. Sci. U.S.A.* **89**, 8317–8321 (1992).
- Kawabata, S.-I., Nakagawa, K., Mita, T., Iwanaga, S. & Davie, E. W. *J. biol. Chem.* **268**, 12498–12503 (1993).
- Mohler, K. M. *et al. Nature* **370**, 218–220 (1994).

FIG. 1 The left column shows wild type, and the right, *7 bcd*. a, Dissected stage-6 embryos stained for *even-skipped* in black and *fushi tarazu* in brown⁹, arrows mark the positions of the cephalic furrows. b, Detail of a; note that the parasegments are more crowded in *7 bcd*; parasegment 8 has the fewest cells and is outlined by small dots. c, Embryos at stage 9 stained for *engrailed* protein. Note the larger head in the *7 bcd* embryo and the differing arrangements of the parasegments (numbered). d, A3 and A4 denticle belts of the first-stage larvae; note the depauperate pattern of denticles in the A3 belt of the *7 bcd* larva. e, A3 of third-stage larvae carrying a *lacZ* enhancer trap line SW005 that marks all the nuclei²⁹; it is apparent that the cell number in the *7 bcd* strain is reduced. f, The nests of histoblasts in the A3 segment of larvae marked by the *lacZ* enhancer trap l(2)4B7 (ref. 30). In these examples the AD, PD and V (anterior dorsal, posterior dorsal and ventral nests) contain 17, 6 and 12 cells in the wild type and 12, 2 and 7 cells in the *7 bcd* larvae. g, Adult A3 dorsal cuticle: 132 bristles can be counted in the wild type and 99 in the *7 bcd*. 2,500 embryos from *7 bcd* females (*P[bcd]/P[bcd]*; *P[bcd]/P[bcd]*; *P[bcd]/TM3 Sb*) were followed; after 24 h the embryos were checked for hatching and for mouth hooks: 45% did not show mouth hooks (although they might have secreted some cuticle), 23% had mouth hooks but did not hatch, 24% hatched, and 8% gave adults. The ventral denticle belts of 63 hatched larvae were studied: 0% had defects in T1 or T2, 1% in T3, 15% in A1, 25% in A2, 85% in A3, 70% in A4, 65% in A5, 55% in A6, 45% in A7 and 35% in A8. The abdomina of 128 adult flies were studied, 55% showed gross defects such as fused segments, the others looked fairly normal, but had reduced bristle numbers in A3. To estimate the number of cells in the embryo, the dorsal and ventral limits were chosen to include all the epidermal progenitors and all the cells that were bounded by the anterior limits of two *fushi tarazu* stripes were counted; this amounted to two parasegments⁹. In *7 bcd* embryos at stage 6, the reduction in number when compared to wild type was 10% for parasegments 2+3, 15% for 4+5, 18% for 6+7, 31% for 8+9 and 21% for 10+11. The numbers of cells in the A3 segment of third-stage larvae were counted. We find 567 ± 20 (s.d.) in wild types and 372 ± 21 for the A3 segment in *7 bcd* larvae, corresponding to a reduction of 35% compared with wild type. The histoblasts were counted: In the A3 wild-type segment, the AD nest has 16 ± 2 , the PD 6 ± 1 and the V 12 ± 3 ; this is similar to previous data¹⁵. In the *7 bcd* larvae, the nests contain 10 ± 2 , 3 ± 2 and 8 ± 2 respectively, indicating a reduction in histoblasts of about 40%. The adult bristles were counted from mounted abdomina, the wild-type male has 124 ± 5 and the *7 bcd* has 91 ± 5 , a reduction of 25%. (Note that the patterns of bristles and pigment are imperfect). Does the number of bristles correlate with the number of cells? Photographs of the bristly area of A3 of newly emerged adults of SW005 and *7 bcd* were taken, stained for β -galactosidase, sample areas were defined; a similar number of trichomes per bristle and a similar number of trichomes per cell were found in the wild type (8.7 and 7.6 ± 0.7 , respectively) and in *7 bcd* (8.1 and 6.9 ± 0.7). We therefore conclude that the bristle number correlates with cell number. SW005 embryos 1.5–3.5 h old were irradiated with 1,000 r; after 5 days of development only 5% survived, these were stained for β -galactosidase; in A3 there were 466 ± 45 cells corresponding to an 18% reduction.



late cleavage and at the blastoderm stage. We thought that random cell death might lead to repair and regrowth, as has been observed for the imaginal discs^{16,17}. Although we cannot rule out the possibility of some regulation of the embryonic epidermis following irradiation, it is clear that there is no recovery to wild type, for these larvae have about 20% fewer cells in their abdominal segments (Fig. 1, legend).

Our findings are surprising because earlier studies suggested that regulation of cell number in *Drosophila* is part of normal development: embryos irradiated with ultraviolet light can compensate for loss of nuclei by extra divisions¹⁸. Also, it has long been known that the imaginal discs undergo extra cell divisions to compensate for cell loss¹⁹. For example, it has been estimated that irradiation (1,000 r) kills some 50% of the cells in the wing disc but the adult structures are normal¹⁷. Further, clones of cells that grow relatively fast compete out other cells in the same compartments²⁰, yet the resulting compartments are normal in cell number and pattern. Finally, when abnormally small segments are made by grafting experiments in the hemimetabolous insect *Oncopeltus*, the segments regulate towards normal size in subsequent moult cycles²¹.

Some failure of regulation had been noted previously, however; if embryos are damaged by ultraviolet irradiation or larvae are cauterized, the defects usually persist^{22,23}—but this damage normally includes segment borders that might be required for the organization of pattern. Our present studies of *7bcd* have the advantage that there is no physical intervention, only the numbers of cells of the primordia are altered and we are able to count the cells directly, both when the primordia are founded and following cell divisions in the embryo. We do not know whether our results rule out the regulation of cell number in the abdomen in all circumstances. But they do favour the hypothesis that the cuticle phenotypes of segment polarity and pair-rule mutations²⁴ result not from cell death and regeneration but from a more direct reorganization of pattern—a reorganization that could be due to the addition or elimination of compartment boundaries in the mutants²⁵.

We have also questioned whether downregulation of cell number occurs in *oskar* mutant embryos²⁶, in which parasegment 5 at gastrulation contains about twice the number of cells as parasegment 4. Measuring the distances between Keilins' organs in the cuticle of the fully developed embryos shows that parasegment 5 is still twice the size of parasegment 4, suggesting that there is no significant downregulation of cell number (data not shown).

It could be that our results reflect the special character of dipteran development: in the brief embryonic stage there may be insufficient time to allow for regulation. Later, at metamorphosis, it is thought that the adult pattern is not built *de novo* but on a template made of the larval pattern; the larval cells are supplanted, area by area, by the proliferating and migrating adult cells, with the larval segment boundaries being the last landmarks to be replaced^{15,27}. Any defects in pattern or size in the template are therefore passed on to the adult abdomen—and this could explain our finding that the adult A3 segment does not regulate. Finally, competition between cells does not occur during growth of the histoblasts but is found in the growing imaginal discs²⁸ where regulation of both cell number and pattern is well established. It is possible that the abilities of cells to compete with each other and to regulate cell number are both products of one mechanism. □

Received 24 March; accepted 14 July 1994.

1. Wigglesworth, V. B. *The Principles of Insect Physiology* p 61 (Methuen, London, 1965).
2. Lawrence, P. A. *The Making of a Fly. The Genetics of Animal Design* (Blackwell Scientific, Oxford, 1992).
3. Lawrence, P. A. *Developmental Systems: Insects* Vol. 2 (eds Counce, S. J. & Waddington, C. H.) 157–209 (Academic, London, 1973).
4. Becker, H. J. Z. *indukt Abstamm. Vererblehre* **88**, 333–373 (1957).
5. Garcia-Bellido, A. & Merriam, J. R. *Dev Biol.* **26**, 264–276 (1971).
6. Berleth, T. et al. *EMBO J.* **7**, 1749–1756 (1988).
7. Driever, W. & Nüsslein-Volhard, C. *Cell* **54**, 95–104 (1988).

8. Nüsslein-Volhard, C. & Roth, S. in *Cellular Basis of Morphogenesis* (eds Evered, D. & Marsh, J.) **144**, 37–91 (Ciba Symp. Fdn, London, 1989).
9. Lawrence, P. A. & Johnston, P. *Development* **105**, 761–767 (1989).
10. Vincent, J.-P. & O'Farrell, P. H. *Cell* **68**, 923–931 (1992).
11. Campos Ortega, J. A. & Hartenstein, V. *The Embryonic Development of Drosophila melanogaster* (Springer, Berlin and Heidelberg, 1985).
12. Roseland, C. R. & Schneiderman, H. A. *Roux's Arch. dev. Biol.* **186**, 235–265 (1979).
13. Driever, W. in *The Development of Drosophila melanogaster* Vol. 1 (eds Bate, M. & Martinez-Arias, A.) (Cold Spring Harbor Laboratory Press, 1994).
14. Lawrence, P. A., Johnston, P., Macdonald, P. & Struhl, G. *Nature* **328**, 440–442 (1987).
15. Madhavan, M. M. & Madhavan, K. J. *Embryol. exp. Morph.* **60**, 1–31 (1980).
16. Schweizer, P. *Biophysik* **8**, 158–188 (1972).
17. Haynie, J. L. & Bryant, P. J. *Roux's Arch. dev. Biol.* **183**, 85–100 (1977).
18. Yasuda, G. K., Baker, J. & Schubiger, G. *Genes Dev.* **5**, 1800–1812 (1991).
19. Bryant, P. J. *J. exp. Zool.* **193**, 49–78 (1975).
20. Simpson, P. & Morata, G. *Dev Biol.* **85**, 299–308 (1981).
21. Wright, D. A. & Lawrence, P. A. *Dev Biol.* **85**, 317–327 (1981).
22. Santamaria, P. & Garcia-Bellido, A. J. *Embryol. exp. Morph.* **28**, 397–417 (1972).
23. Lohs-Schardin, M., Cremer, C. & Nüsslein-Volhard, C. *Dev Biol.* **73**, 239–255 (1979).
24. Nüsslein-Volhard, C. & Wieschaus, E. *Nature* **287**, 795–801 (1980).
25. Sampedro, J., Johnston, P. & Lawrence, P. A. *Development* **117**, 677–687 (1993).
26. Lehmann, R. & Nüsslein-Volhard, C. *Cell* **47**, 141–152 (1983).
27. Madhavan, M. M. & Madhavan, K. J. *Embryol. exp. Morph.* **82**, 1–8 (1984).
28. Morata, G. & Ripoll, P. *Dev Biol.* **42**, 211–221 (1975).
29. Meise, M. & Janning, W. *Development* **118**, 1107–1121 (1993).
30. Hartenstein, V. & Jan, Y. N. *Roux's Arch. dev. Biol.* **194**, 308–325 (1992).

ACKNOWLEDGEMENTS. We thank C. Nüsslein-Volhard who suggested this project years ago; G. Struhl for advice; D. St Johnston for the *7bcd* stock; W. Driever, who made it and allowed us to use it before publication; M. Frasch, H. Krause and R. Smith for antibodies; J. Janning and A. Martinez-Arias for stocks; P. Johnston for help and for staining the embryos for eve and ftz; and our colleagues for comments on the manuscript. A.B. was supported by a grant from Biorad.

Archetypal organization of the amphioxus *Hox* gene cluster

Jordi Garcia-Fernández* & Peter W. H. Holland*†

Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK

ORGANIZATION into gene clusters is an essential and diagnostic feature of *Hox* genes¹. Insect and nematode genomes possess single *Hox* gene clusters (split in *Drosophila*); in mammals, there are 38 *Hox* genes in four clusters on different chromosomes^{2,3}. A collinear relationship between chromosomal position, activation time and anterior expression limit of vertebrate *Hox* genes suggests that clustering may be important for precise spatiotemporal gene regulation and hence embryonic patterning^{2,4}. *Hox* genes have a wide phylogenetic distribution within the metazoa, and are implicated in the control of regionalization along the anteroposterior body axis^{2,5}. It has been suggested that changes in *Hox* gene number and genomic organization played a role in metazoan body-plan evolution^{6–8}, but identifying significant changes is difficult because *Hox* gene organization is known from only very few and widely divergent taxa (principally insects, nematodes and vertebrates)³. Here we analyse the complexity and organization of *Hox* genes in a cephalochordate, amphioxus, the taxon thought to be the sister group of the vertebrates⁹. We find that the amphioxus genome has only one *Hox* gene cluster. It has similar genomic organization to the four mammalian *Hox* clusters, and contains homologues of at least the first ten paralogous groups of vertebrate *Hox* genes in a collinear array. Remarkably, this organization is compatible with that inferred for a direct ancestor of the vertebrates; we conclude that amphioxus is a living representative of a critical intermediate stage in *Hox* cluster evolution.

Gene order within *Hox* gene clusters, together with DNA sequence comparison, allows the 38 mammalian *Hox* genes to be divided into 13 paralogous groups; each group comprises genes related by *Hox* cluster duplication^{2,7,10,11}. To investigate

* Present addresses: Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Av. Diagonal 645, 08071 Barcelona, Spain (J.G.-F.); and School of Animal and Microbial Sciences, University of Reading, Whiteknights, PO Box 228, Reading RG6 2AJ, UK (P.W.H.H.).

† To whom correspondence should be addressed.

Antennapedia		RKRGRQTYTRYQTLELEKEPHFNRYLTRRRRIIEAHALCLTERQIKIWFQNRMRKWKKEN		KTKGEPGSGGGEDEITPPNSPQ	
Amphiox-1	KPGEYGTTSK	PNNG-TNF-TK-LT-----Y-K-----A-V-----A-N-N-T-V-----Q--RE	KENGFTPGSGGSPAGEDSPSKST		
m a-1 (92%)	KVGEYGVV.GQ	PNVAV-TNF-TK-LT-----K-A-A-V-----AS-Q-N-T-V-----Q--RE	KEGLLPISPATPGSDEKTEESSEKSSPSPSPSPASSTSDLTSTSH		
m b-1 (87%)	AKVSELGL..GP	PGGL-TNF-TR-LT-----K-S-A-V-----AT-E-N-T-V-----Q--RE	REGGRMPAGPPGCPKAAAGDASDQACTSPFASPSSITS		
m d-1 (83%)	KLSEYGAT.SP	PSAI-TNFSTK-LT-----K-A-----NC-Q-NDT-V-----Q--RE	REGLLATAASVASIKLPRSETSPKSGRNLGSPSQAQEPS		
d lab (82%)	...SGLSSCSSLSSNT	NNS--TNF-NK-LT-----A-----NT-Q-N-T-V-----Q--RV	KEGLIPADILTQHSTSVISEKPPQQQPPELQLKSGGSLDGLNELATGAPS...		
Amphiox-2	NADFLTSPDQVNS	SR-L-TVF-NT-L-----Y-K-VCKP-K--SY-D-N--V-----RQ-RRD	TKGRSEIGTDPGTAAPADGGRAEEGSGGGSAACFRPQG		
m a-2 (77%)	SLEIADGSGGG	SR-L-TA-NT-L-----K-C-P-V--AL-D--V--V-----H-RQT	QKENQNSSEGRKPNLEDSKVEEDEEEKSLFEQALSVSGALLEREYTFQON...		
m b-2 (77%)	GFGLPECGSGG	SR-L-TA-NT-L-----K-C-P-V--AL-D--V--V-----H-RQT	E...		
d pb (75%)	EFVPENGL	PR-L-TA-NT-L-----K-C-P-----AS-D--V--V-----H-RQT	LSKTDDHDKDSLKGDDQSDSNSNKKSCQGCLEPDDIPDSTSNRSHNNL...		
Amphiox-3	GTTEPGESPLGGAA	G-A-TA--SA-LV-----C-P-V-M-AM-N-----Y--Q	KVKGSGSGGSGGMSNPSPPATTTTPGINPGLPHPTTQPSLSQSNMNSNHM...		
m a-3 (87%)	...GPPGQAS	S--TA--P-LV-----M-P-V-M-NL-N-----Y--DQ	KGKGLM...		
m b-3 (92%)	EGCG ₁₆ SSG ₁₀ DKSPGSA	S-A-TA--SA-LV-----C-P-V-M-NL-N-S-----Y--DQ	KAKGLASSSGGSPAGSPFPQFMQSTAGFMNALHMSPTSYDPSPPAFGKHQ...		
m d-3 (90%)	ENECDEKSPGP.A	S-V-T--SA-LV-----C-P-V-M-NL-N-----Y--DQ	KAKGLHSPAGQSPERSFPLGGAAGHVAYSGQLVPVGLAYDAPSPPAFAKSQ...		
d pb (72%)	EFVPENGL	PR-L-TA-NT-L-----K-C-P-----AS-D--V--V-----H-RQT	LSKTDDHDKDSLKGDDQSDSNSNKKSCQGCLEPDDIPDSTSNRSHNNL...		
Amphiox-4	STSYNGQD	T-S-TA--Q-V-----S-G-----D-	RLPNTKTRSSASGGTGGQRFSSTTQPSRGLSGVCL		
m a-4 (92%)	VNSSYNGGE	P-S-TA--Q-V-----T--S--V-----DH	KLPTNKMRSSNTASAPAGPPKQATHSPHHHPHLPAGSTPIPSII		
m b-4 (88%)	VNPNYAGGE	P-S-TA--Q-V-----Y-----V-----S-----DH	KLPTNKIRSGGTAGAGGPPGRFNGGPPAL		
m c-4 (92%)	VNPNYNGGE	P-S-TA--Q-V-----Y-----S-----DH	RLPNTKVRSAAPPAGAAFPSTLSAATPGTSEDHQSASATPEQQRAEDITRL		
m d-4 (92%)	ANPNYTGGE	P-S-TA--Q-V-----T--P-----DH	KLPTNKGRSSSSSSCASSAAPGQHLQPMADHHTDLTTL		
d DId (87%)	NGSYQFGME	P-Q-TA--H-I-----Y-----T-V-S-----D-	KLPTNKVRKKTVDANGKPTPAKPTKRAASKKQQAQQQQTQQTQQTQVPM...		
Amphiox-5	AGTGD	N--T-TA-----	KLKSLSGCQQTQGLNPN		
m a-5 (93%)	DNTGGPE	G-A-TA-----S-----D-	KLKSMMAAGGAFRP		
m b-5 (93%)	DMTGPD	G-A-TA-----S-----D-	KLKSMSLATAGSAFP		
m c-5 (88%)	ETD	G-S-TS-----NN--N-----DS	KMKSEAL		
d Scr (93%)	TVNANGE	T-Q-TS-----H	KMASMNIVPYHMGYPHQFDIHPSQFAHLSA		
Amphiox-6	MGEF	K-----K-----L-G-----	KIPSLNATTINQLTMNHNHNGVNSDTDKERDIK		
m a-6 (88%)	AVYGS	GR-----N-----S	KLINSTNAGDSEAK		
m b-6 (87%)	SSFGPS	GR-----Y-----S	KLKSLASQLSAEEEEKPAE		
m c-6 (83%)	VGVGAD	-R-----I-S-----N-----S	NLTSTLGGGGGATADSLGGKEEKREEKEEKKEKQE		
d Ubx (83%)	TNGL	-R-----T-H-----M-----L-L-L	QAIKELNEQEKQQAQKAAAAAAA/VQGGHLDQ		
d Abd-A (88%)	PNGCP	-----F-----L	RAVKEINEQARRDREEQEKMAQETMKSAQNKQVQQQQQQQQQQQQQQQ...		
Amphiox-7	PE	-----K-----	KLKSLKQPAESETSTSTS		
m a-7 (97%)	PD	-----H	KDESQAPTAAPEDAVPVSSTAADKADEEEEEEEEEEEEEEEE		
m b-7 (95%)	PD	-----Y-----T-----	KTSGGPTTGQDKAEGEEEE		
Amphiox-8	PE	-R-----K-----G-----L--A	AMLCPKPAETETEKSSDQSEKSEEV		
m b-8 (87%)	AG	-R-----S-----L-P-K-----VS-----G-----V-----	NKDKFSSSKCEQEELEKEKLERAPETAEGDAQKDKK		
m c-8 (85%)	PG	-RS-----L-P-K-----VS-----G-----V-----	NKDKLPGARDEEKVEEGNEEEKEEKEENKD		
m d-8 (82%)	APG	-R-----S-F-----L-P-K-----VS-T-A-----V-----	NKDKFPASRPDGKDGDPKKEASGLEEDGADGCPIN		
Amphiox-9	SUVANQPGWMNHS	SRKK-CP--F-----LY-M--E-Y--SQHVN--V-----M--MS	KQRQEQQQSPQ		
m a-9 (80%)	NPAANWLHARS	TRKK-CP--KH-----L-M--D-Y-V-RL-N--V-----M--I	KDRAKDE		
m b-9 (80%)	NFSANWLHARS	SRKK-CP--K-----L-M--D-H-V-RL-N-S--V-----M--M	NGAGKE		
m c-9 (82%)	NPVANWLHARS	TRKK-CP--K-----L-M--D-Y-V-RV-N--V-----M--M	KEKTKEQS		
m d-9 (83%)	NPAANWLHARS	TRKK-CP--K-----L-M--D-Y-V-RI-N--V-----M--MS	KEKCPKD		
d Abd-B (67%)	...VGPCTFNPLHEW	VRKK-KP-SKF-----L-A-VSKQK-W-L-RN-Q--V-----N--NS	QRQANQQNNNNSSNNHQAQATQQRQANQQHSGHHLNLSLNMGHAAKMHQ		
Amphiox-10	SCGATSWMAPRV	GRKK-CP--K--I-----L-M-VS-E-Q--SRHVN-SD--V-----M-RM	KAREEQIRNQQRGHQAVPVSG		
m a-10 (83%)	SKGENAANWLAKS	GRKK-CP--KH-----L-M--E-L--SRSVH--D--V-----L-M	..RENRIRELTANFNPS		
m c-10 (82%)	EIKAENTTGNWLAKS	GRKK-CP--KH-----L-M--E-L--SKTIN--D--V-----L-M	..RENRIRELTNSNFT		
m d-10 (82%)	EIKSDTPTSNWLAKS	GREK-CP--KH-----L-M--E-L--SKSVN--D--V-----L-MS	..RENRIRELTANLTPS		
d Abd-B (65%)	...VGPCTFNPLHEW	VRKK-KP-SKF-----L-A-VSKQK-W-L-RN-Q--V-----N--NS	QRQANQQNNNNSSNNHQAQATQQRQANQQHSGHHLNLSLNMGHAAKMHQ		

FIG. 1 Alignments between deduced homeodomain sequences of amphioxus *Hox* genes and those encoded by mouse *Hox* and *Drosophila* HOM-C genes^{24,25}. Dashes indicate amino-acid identity to Antp in the homeodomain only; dots indicate gaps introduced to optimize alignments. Also shown are the complete C-terminal coding regions for nine of the ten amphioxus genes (*Amphiox-3* is from ref. 13) aligned with these regions of mouse *Hox* and *Drosophila* HOM-C proteins (except where very long or, for two mouse genes, not available, indicated by a

dotted line); plus residues N-terminal of the homeodomain up to the deduced intron (except for *Drosophila* lab and Abd-B, where the intron is much further upstream, indicated by a dotted line). Flanking sequences are from the EMBL database and references therein. Percent identities, with respect to *Amphiox* homeodomains, cover the homeodomain only. Sequences have been deposited with the EMBL/GenBank/DBJ databases and assigned the accession numbers Z35142–Z35150.

the origin of the four clusters and the paralogous groups, we examined the organization of *Hox* genes in a species of amphioxus, *Branchiostoma floridae*. Using the polymerase chain reaction (PCR), we initially cloned short fragments from five amphioxus *Hox* genes¹²; here we use several primer sets complementary to *Hox* genes from paralogous groups 1–10 (not expected to amplify genes from groups 11–13) to identify nine *Hox* genes whose DNA sequences were compatible with, but not proof of, a single cluster. To obtain more extensive sequences, we screened an amphioxus larval complementary DNA library with a heterogeneous *Hox* gene PCR band. This yielded partial cDNA clones from four of the amphioxus *Hox* genes. A representative genomic library was also constructed from adult amphioxus DNA and screened in turn with eight of the nine classes of PCR clones. Reduced stringency conditions were chosen to allow cross-hybridization between related *Hox* sequences, enabling most, if not all, amphioxus *Hox* genes related to paralogous groups 1 to 10 to be detected. Multiple

clones from each screening were analysed, restriction-endonuclease mapped, and the homeobox regions sequenced. All nine genes identified by PCR were represented in the genomic clones (confirming their authenticity), plus one additional *Hox* gene.

Comparison of complete homeobox sequences suggests that each gene can be assigned to a different paralogous group of mammalian *Hox* genes, representing every paralogous group from 1 to 10 (Fig. 1). Accordingly, we designate these genes *Amphiox-1* to *Amphiox-10* (of these, the *Amphiox-3* gene has been reported previously¹³). Assignment of *Amphiox-1* to *Amphiox-5* to each paralogous group from 1–5 is unambiguous and based on multiple uniquely shared amino-acid residues. *Amphiox-9* and *Amphiox-10* share clear similarities with both paralogous groups 9 and 10, but we noted shared residues C-terminal of the homeodomain (QS for group 9, IR for group 10) and conserved residues in the homeodomain in group 9 genes (Y at position 32) or group 10 genes (G at position 1, D at position 42). Although the *Amphiox-8* homeodomain lacks

FIG. 2 Alignment between complete deduced protein sequence of *AmphiHox-6* and mouse *Hoxb-6* (ref. 26). Asterisks indicate identities; dots indicate gaps introduced to optimize alignment; arrowheads indicate intron positions. Database searches (TFASTA), using either the complete coding sequence, or the first exon, revealed *Hoxb-6* as the most similar protein.

```

AmphiHox-6  MSSYFVNS.FAGNLQNGGDTYPAGQG..YPVRY.DPMRQYPTYPGPPERF
Mouse b-6    MSSYFVNSTFPVTLASGQESFL.GQLPLYSSGYADPLRHPAPYG.PGPG
            ***** * * * * * * * * * * * * * * * *

AmphiHox-6  PVYSTGDDGLYGAQQGAY..TDPIKHEAALASHHLPQYSSCKLRNTNTSP
Mouse b-6    QDKGFAASSYPPAGGGYGRAAPCDYGPAPAFYREKDAACALSGADEPPP
            * * * * * * * * * * * * * * * *

AmphiHox-6  LPSAQGSPGST.....VGAQLPQPTPPVFWMRKGSQTA...MGEKKR
Mouse b-6    FHPEPRKSDCAQDKSVFGETEQQKSTPVYPMQRMNSCNSSSFGPSGRR
            * * * * * * * * * * * * * * * *

AmphiHox-6  GRQTYTRYQTLEKEFHFNKYLTRKRRIEIAHLGLTERQIKIWFQNR
Mouse b-6    GRQTYTRYQTLEKEFHFNKYLTRKRRIEIAHLGLTERQIKIWFQNR
            ***** * * * * * * * * * * * * * * * *

AmphiHox-6  MKWKKENKIPSLNATTINQLTNMNSHNGVNSDTERDKID
Mouse b-6    MKWKKESKLLSASQLSAEEEEKPAE.....
            ***** * *

```

several unique amino-acid residues shared by the three mouse group-8 genes, we are confident of this assignment because of a unique serine at position 9, glycine at position 39, and additional shared residues C-terminal of the homeodomain.

AmphiHox-6 and *AmphiHox-7* are more problematic: the genes encode very similar homeodomain sequences and there are no uniquely diagnostic residues shared by all mammalian group 6 or group 7 homeodomains. Nonetheless, it is striking that the *AmphiHox-7* homeodomain has 58 and 57 identities (from 60 amino acids) to mouse *Hoxa-7* and *Hoxb-7*, respectively; these are the highest sequence identities shown by any of the amphioxus-mouse comparisons. This evidence for assignment to group 7 is also reinforced by identity with mouse *Hoxa-7* at four positions C-terminal of the homeodomain. In contrast,

comparison of homeodomains and C-terminal regions does not reveal whether *AmphiHox-6* is a direct homologue of the group 6 genes of mammals or, alternatively, is derived from an independent gene-duplication event. To investigate further, we determined the complete coding sequence for the *AmphiHox-6* gene from cDNA and genomic clones. Sequence alignment (Fig. 2) reveals extensive amino-acid identity with mouse *Hoxb-6* throughout the deduced protein, confirming its assignment to this paralogous group.

These assignments reveal that amphioxus, in contrast to *Drosophila* and nematodes^{2,3}, but like mammals and birds^{2,4,14}, has more than one gene related to *Drosophila Abd-B* and has representatives of paralogous groups 6, 7 and 8 (the *Antp/Ubx/abd-A*-related genes). Only one representative of each group was

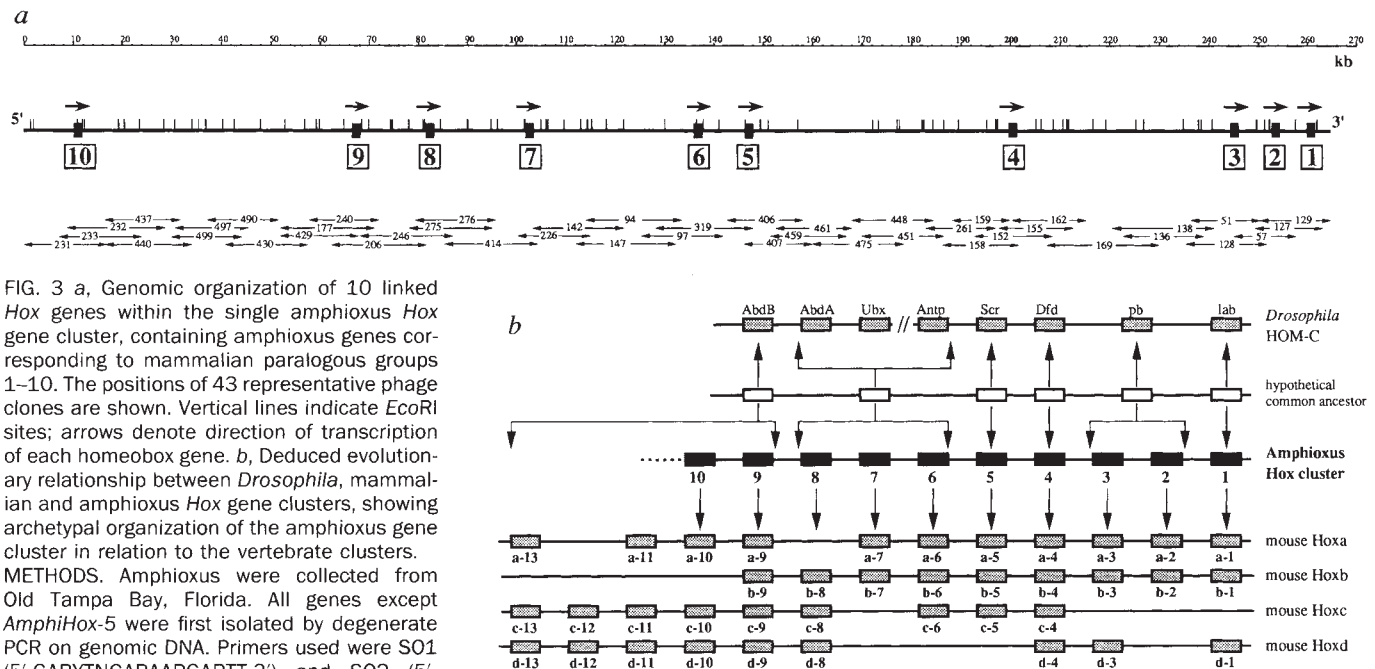


FIG. 3 a, Genomic organization of 10 linked *Hox* genes within the single amphioxus *Hox* gene cluster, containing amphioxus genes corresponding to mammalian paralogous groups 1–10. The positions of 43 representative phage clones are shown. Vertical lines indicate *EcoRI* sites; arrows denote direction of transcription of each homeobox gene. b, Deduced evolutionary relationship between *Drosophila*, mammalian and amphioxus *Hox* gene clusters, showing archetypal organization of the amphioxus gene cluster in relation to the vertebrate clusters. METHODS. Amphioxus were collected from Old Tampa Bay, Florida. All genes except *AmphiHox-5* were first isolated by degenerate PCR on genomic DNA. Primers used were S01 (5'-GARYTNGARAARGARTT-3') and S02 (5'-CKNCKRTTYTGRAACCA-3'), corresponding to homeodomain residues 15–20 and 48–53 of groups 1–10 *Hox* genes. Cycling conditions and cloning strategies were as described²⁷, with annealing temperatures of 37, 45 or 52 °C. The nine genes were represented in 35 clones; sequencing of 93 additional clones from PCR amplification of genomic and cDNA libraries yielded no additional *Hox* genes. About 5×10^5 clones of a *B. floridae* cDNA library in λ ZAP (Stratagene) were screened at moderate stringency (46% formamide, 40 °C) with a 116-bp heterogeneous *Hox* gene PCR band, pooled from different PCRs. Multiple clones of *AmphiHox-1*, -4, -6 and -9 were isolated; the homeobox sequences exactly matched the PCR and genomic clones. To construct the genomic library, high-molecular-weight amphioxus DNA was partially digested with *Sau3A*, size-fractionated on a sucrose gradient (15–25 kb), and cloned into λ FIX II (Stratagene). About 5×10^5 recombinant clones were screened sequentially with probes from PCR clones of *AmphiHox-1*, -2, -3, -4, -7, -8, -9 and -10, under moderate

stringency conditions (hybridization in 50% formamide at 37–40 °C; washes $2 \times$ SSC, 0.1% SDS, 55 °C). Positive clones were restriction-mapped, hybridizing bands cloned into plasmid vectors and homeoboxes and flanking regions sequenced. The *AmphiHox-1*, -2, -3, -4, -9 and -10 probes resulted in isolation of one genomic region in each case; screenings with *AmphiHox-7* and -8 probes produced clones of *AmphiHox-4*, -5, -6, -7 and -8, confirming that the hybridization conditions used could detect related genes. For chromosome walking, the terminal *EcoRI*/*NotI* or *HindIII*/*NotI* ends of selected lambda clones were used as probes to isolate adjacent genomic regions. All *Hox* genes independently detected belong to a single *Hox* cluster. To confirm the absence of other homeoboxes, all genomic regions were hybridized with the oligonucleotide cHB-1 (5'-CKNCKRTTYTGRAACCADATYTT-3')²⁸ under conditions allowing detection of genes with up to two nucleotide mismatches from the *Hox*-type homeodomain helix III (ref. 29).

obtained, indicating that a single set of *Hox* genes is present in the amphioxus genome.

We used a genomic walking analysis comprising ~200 overlapping phage clones to investigate whether the genes were chromosomally linked. This revealed that the 10 genes (*AmphiHox-1* to *AmphiHox-10*) are linked in the genome, in a collinear array (5' *AmphiHox-10* to 3' *AmphiHox-1*) spanning a region of 270 kilobases (kb) (Fig. 3a). Position in the cluster is consistent with the sequence comparisons used to ascribe each *AmphiHox* gene to a particular paralogous group (Fig. 1). We did not extend the walk 5' of *AmphiHox-10*, where homologues of the more divergent group 11–13 genes may be present.

The amphioxus *Hox* genes are similar in structure to their vertebrate counterparts: the deduced direction of transcription is the same for all genes, the homeobox is situated in the last exon, and an intron splice site is present just upstream of the homeobox. Intergenic distances are similar to those in mammalian *Hox* clusters^{15,16} or slightly larger (for example, *AmphiHox-3* to *AmphiHox-4* is 45 kb, *AmphiHox-4* to *AmphiHox-5* is 55 kb, and *AmphiHox-9* to *AmphiHox-10* is 58 kb). Southern hybridization was used to confirm the absence of other *Hox* genes in the intergenic regions. No particular mammalian *Hox* cluster shows consistently higher sequence similarity to the amphioxus *Hox* genes, indicating that the four mammalian *Hox* clusters are equally divergent with respect to amphioxus (Fig. 1). Our results demonstrate that amphioxus has a single *Hox* gene cluster of remarkably vertebrate-like organization.

Amphioxus has been included in an extensive PCR-based phylogenetic survey of *Hox* genes¹⁷, which postulated the presence of two *Hox* gene clusters. This claim contradicts our genomic cloning results, so we compared our complete homeobox sequences with the PCR-derived clones of the earlier study and found that some of the latter had been misassigned to paralogous groups. This highlights a general problem when using short PCR clones to investigate *Hox* gene complexity. Also, we could not detect two of the clones central to the two-cluster model¹⁷ (second clones from groups 1 and 3) by PCR, genomic cloning or cDNA cloning. To investigate the origin of these clones, we designed complementary oligonucleotides to both types of group 1 and group 3 genes, and used them as gene-specific primers in PCR amplification of amphioxus genomic DNA. Positive controls of plasmids engineered to contain the target sequences were included. In addition, the four oligonucleotides were used as probes to screen Southern blots of heterogeneous *Hox* gene PCR products. Both approaches showed that the two additional clones reported earlier¹⁷ were not derived from the amphioxus genome (data not shown); we suggest that they may derive from an extraneous organism (perhaps a parasite or food item of amphioxus).

The amphioxus *Hox* gene cluster has an unprecedented organization of particular interest from an evolutionary perspective. Mammals, birds, amphibia, fish and probably lampreys have multiple *Hox* clusters^{2,7,14,17–20}, thought to have arisen by gene-

cluster duplication close to vertebrate origins^{7,8,10,11}. Phylogenetic reconstructions suggest that immediately before it was duplicated, this ancestral vertebrate *Hox* cluster contained either nine or thirteen genes (depending on whether *Abd-B* tandem duplications occurred before or after *Hox* cluster duplication)^{4,10}. Subsequent to *Hox* cluster duplication, genes have been secondarily lost; none of the mammalian clusters contains genes from all paralogous groups. One of the most intriguing features of the amphioxus *Hox* gene cluster is the similarity between its organization and that inferred for the last direct ancestor of the vertebrates to possess a single *Hox* cluster. The proposed one-to-one correspondence between amphioxus *Hox* genes and vertebrate *Hox* paralogous groups implies that there have been no additional tandem duplications, or secondary losses, of *Hox* genes in the lineage leading to amphioxus, despite the fact that the cephalochordate and vertebrate lineages have been separate for over 520 million years²¹. The amphioxus genome has apparently retained the ancestral and archetypal pre-vertebrate, pre-duplication *Hox* cluster organization (Fig. 3b).

The results demonstrate that tandem gene duplication giving at least two *Abd-B*-related genes (groups 9 and 10) occurred before *Hox* cluster duplication¹⁰. Several authors have suggested links between the multiple vertebrate *Abd-B*-related genes and the evolution of either posterior specializations⁴, genital ducts⁸ or paired appendages^{8,22}. The amphioxus results imply that the initial selective pressures for retention of tandem *Abd-B*-related genes cannot have been the evolution of features unique to vertebrates.

Hox genes are involved in patterning the body plan, hence the archetypal organization of the amphioxus *Hox* cluster presents an intriguing parallel to the old idea that the anatomy of amphioxus is similar to that of the immediate ancestor of the vertebrates⁹. The results also imply that *Hox* cluster duplications occurred after the divergence of the cephalochordate and vertebrate lineages, close to vertebrate origins. Human and mouse *Hox* clusters form part of more extensive paralogy groups of related linked genes on different chromosomes²³, hence *Hox* cluster duplication coincided with the enlargement of several gene families. We suggest that the origin of new genes played an important permissive role in the evolution of complex vertebrate body plans. Finally, there are practical implications. Investigation of *Hox* gene functions in mammals, for example by gene targeting, is hampered by the presence of multiple gene clusters and partial genetic redundancy^{2,8}. Functional analyses in standard model invertebrate species (*Drosophila*, nematodes) is simpler owing to the possession of a single *Hox* cluster, but extrapolation to vertebrates is hampered by their extreme phylogenetic distance, dissimilar body plans and divergent *Hox* cluster organizations. Amphioxus has a chordate body plan, a close phylogenetic affinity to vertebrates and a single, vertebrate-like, *Hox* cluster; this combination will prove particularly useful in studies of the regulation and function of chordate *Hox* genes. □

Received 6 April; accepted 12 July 1994.

1. Scott, M. *Cell* **71**, 551–553 (1992).
2. McGinnis, W. & Krumlauf, R. *Cell* **68**, 283–302 (1992).
3. Bürglin, T. & Ruvkun, G. *Curr. Opin. Gen. Dev.* **3**, 615–620 (1993).
4. Izpisua-Belmonte, J.-C. et al. *EMBO J.* **10**, 2279–2289 (1991).
5. Slack, J. M. W., Holland, P. W. H. & Graham, C. F. *Nature* **361**, 490–492 (1993).
6. Akam, M. et al. *Development* **104**, suppl., 123–133 (1988).
7. Kappen, C., Schughart, K. & Ruddle, F. H. *Proc. natn. Acad. Sci. U.S.A.* **86**, 5459–5463 (1989).
8. Holland, P. W. H. *BioEssays* **14**, 267–273 (1992).
9. Willey, A. *Amphioxus and the Ancestry of the Vertebrates* (Macmillan, New York, 1894).
10. Kappen, C. & Ruddle, F. H. *Curr. Opin. Gen. Dev.* **3**, 931–938 (1993).
11. Schubert, F. R., Nieselt-Struwe, K. & Gruss, P. *Proc. natn. Acad. Sci. U.S.A.* **90**, 143–147 (1993).
12. Holland, P. W. H., Garcia-Fernández, J., Holland, L. Z., Williams, N. A. & Holland, N. D. *J. Mar. Biol. Assoc.* **74**, 49–60 (1994).
13. Holland, P. W. H., Holland, L. Z., Williams, N. A. & Holland, N. D. *Development* **116**, 653–661 (1992).
14. Izpisua-Belmonte, J.-C. et al. *Nature* **350**, 585–589 (1991).
15. Acampora, D. et al. *Nucleic Acids Res.* **17**, 10385–10402 (1989).
16. Rubock, M. J. et al. *Proc. natn. Acad. Sci. U.S.A.* **87**, 4751–4755 (1990).
17. Pendetton, J. W., Nagai, B. K., Murtha, M. T. & Ruddle, F. H. *Proc. natn. Acad. Sci. U.S.A.* **90**, 6300–6304 (1993).

18. Dekker, E. J. et al. *Mech. Dev.* **40**, 3–12 (1992).
19. Belleville, S. et al. *Gene* **114**, 179–186 (1992).
20. Molven, A., Hordvik, I. & Njolstad, P. R. *Biochim. biophys. Acta* **1173**, 102–106 (1993).
21. Sansom, I. J., Smith, M. P., Armstrong, H. A. & Smith, M. M. *Science* **256**, 1308–1311 (1992).
22. Tabin, C. J. *Development* **116**, 289–296 (1992).
23. Lundin, L. *Genomics* **16**, 1–19 (1993).
24. Bürglin, T. in *A Guidebook for Homeobox Genes* (ed. Duboule, D.) 25–71 (Oxford University Press, Oxford, 1994).
25. Kappen, C., Schughart, K. & Ruddle, F. H. *Genomics* **18**, 54–70 (1993).
26. Schughart, K., Utset, M. F., Awgulewitsch, A. & Ruddle, F. H. *Proc. natn. Acad. Sci. U.S.A.* **85**, 5582–5586 (1988).
27. Holland, P. W. H. in *Essential Developmental Biology: A Practical Approach* (eds Stern, C. D. & Holland, P. W. H.) 243–255 (Oxford University Press, Oxford, 1993).
28. Webster, P. J. & Mansour, T. E. *Mech. Dev.* **38**, 25–32 (1992).
29. Bürglin, T. R., Finney, M., Coulson, A. & Ruvkun, G. *Nature* **341**, 239–243 (1989).

ACKNOWLEDGEMENTS. We thank L. Z. Holland for the amphioxus cDNA library, N. D. Holland and J. Lawrence for assistance during specimen collection, and N. Williams and S. Thompson for laboratory assistance. This research was funded by the SERC, a Royal Society URF to P.W.H.H., and EMBO and Human Frontier Fellowships to J.G.-F.

A conserved retinoic acid response element required for early expression of the homeobox gene *Hoxb-1*

Heather Marshall*, Michèle Studer*, Heike Pöpperl*, Sam Aparicio†, Atsushi Kuroiwa‡, Sydney Brenner† & Robb Krumlauf*

* Laboratory of Developmental Neurobiology, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

† Molecular Genetic Unit and Department of Medicine, University of Cambridge Clinical School, Addenbrooke's Hospital, Cambridge CB2 2QQ, UK

‡ Department of Molecular Biology, School of Science, Nagoya University, Chihusa-Ku, Nagoya, Japan 464-01

WITHIN the *Hoxb* homeobox gene complex, *Hoxb-1* is the earliest member expressed in the mesoderm and neuroectoderm of primitive streak and presomite embryos, preceding rhombomere-restricted expression in the hindbrain^{1,7}. Ectopic exposure of embryos to retinoic acid alters spatial aspects of *Hox* gene expression patterns⁸⁻¹⁵. However, the role of retinoids in regulating these genes during normal development is unclear. We have now identified two enhancers, 3' of the mouse *Hoxb-1* gene, which together reconstruct the early endogenous expression pattern and mediate the early ectopic response to retinoic acid. Furthermore, these regions are functionally conserved in both chicken and pufferfish (*Fugu rubripes*)¹⁶ *Hoxb-1* genes. The enhancer that controls the retinoic acid response, and regulates expression predominantly in neuroectoderm, contains a retinoic acid response element (RARE). Point mutations in the RARE abolish expression in neuroectoderm. Therefore, this RARE is not only involved in the ectopic response to retinoic acid, but is also essential for establishing aspects of the early *Hoxb-1* expression pattern.

A transgenic construct with 18 kilobases (kb) of genomic DNA from the *Hoxb-1* gene can reconstruct both the early and late expression patterns, and its response to exogenous retinoic acid⁹. Deletion analysis in transgenic mice was used to identify the regulatory regions that mediate the early patterns of expression and response to retinoic acid. An internal 7.5-kb *EcoRV* fragment which generates only the late, rhombomere-4-restricted phase of expression (construct 1; Fig. 1a) was used to test the ability of additional 5'- and 3'-flanking regions to restore early *Hoxb-1* expression. Whereas sequences upstream of the 5' *EcoRV* site had no activity, a 3' region (*EcoRV-SalI*; construct 2; Fig. 1a) imposed an early pattern of *Hoxb-1* expression comparable to that of the endogenous gene (Fig. 2a-h)^{1,3,7}. The 3' fragment (*EcoRV-SalI*) on its own was tested on a heterologous promoter (construct 7) to determine whether it is sufficient to regulate the early expression. This fragment functioned as an enhancer, and produced the same temporal and spatial patterns of expression (Fig. 2i,j) observed with the larger transgenic constructs (construct 2; Fig. 1a) and the endogenous gene.

We examined whether this region also mediates the early ectopic response to retinoic acid^{8,9}, and found that retinoic acid rapidly induced expression in the neuroectoderm and mesoderm (construct 7; Fig. 2k-l). Within 4 hours of exposure to retinoic acid at 7.25 days post-coitum (d.p.c.), levels of staining had increased and there was an anterior shift in expression boundaries within the neural plate and mesoderm. Together, these data demonstrate that the 3' *EcoRV-SalI* fragment mediates aspects of the early *Hoxb-1* expression pattern, including the response to retinoic acid.

To investigate whether these elements are conserved in other species, we analysed the chicken and pufferfish¹⁶ *Hoxb-1* genes. At 8.0 d.p.c. a chicken 3' regulatory region (construct 10) directed staining in the node, mesoderm and ectoderm with similar anterior boundaries of expression to those seen in the mouse 3' region (Fig. 2m); however, levels of expression were consistently lower in the ectoderm and higher in the mesoderm. A *Fugu* 3' region (construct 11) also generated the early expression pattern in mesoderm and neuroectoderm, although staining was weak in the posterior mesoderm (Fig. 2o). Both chicken and pufferfish early regulatory regions also responded to retinoic acid, stimulating expression in more anterior domains (Fig. 2n,p). The similarity of the patterns generated by DNA from the three vertebrates implies that the upstream factors and their target sequences, which regulate early *Hoxb-1* expression and response to retinoic acid have been highly conserved.

In the mouse 3' control region, two enhancers have been mapped which regulate different aspects of the early expression. Deletion of a 4.0-kb *HindIII-SalI* fragment (constructs 3 and 9) resulted in a loss of several mesodermal sites of expression, and the *HindIII-SalI* fragment itself (construct 8) functioned as a mesodermal enhancer directing expression in the node and emerging paraxial and lateral plate mesoderm (Fig. 3a-e). The timing of appearance and downregulation of the mesodermal expression were similar to those seen with the endogenous gene. We tested whether this enhancer mediates the *Hoxb-1* mesodermal response^{8,10} to retinoic acid, and found that expression did not change on exposure to retinoic acid (data not shown).

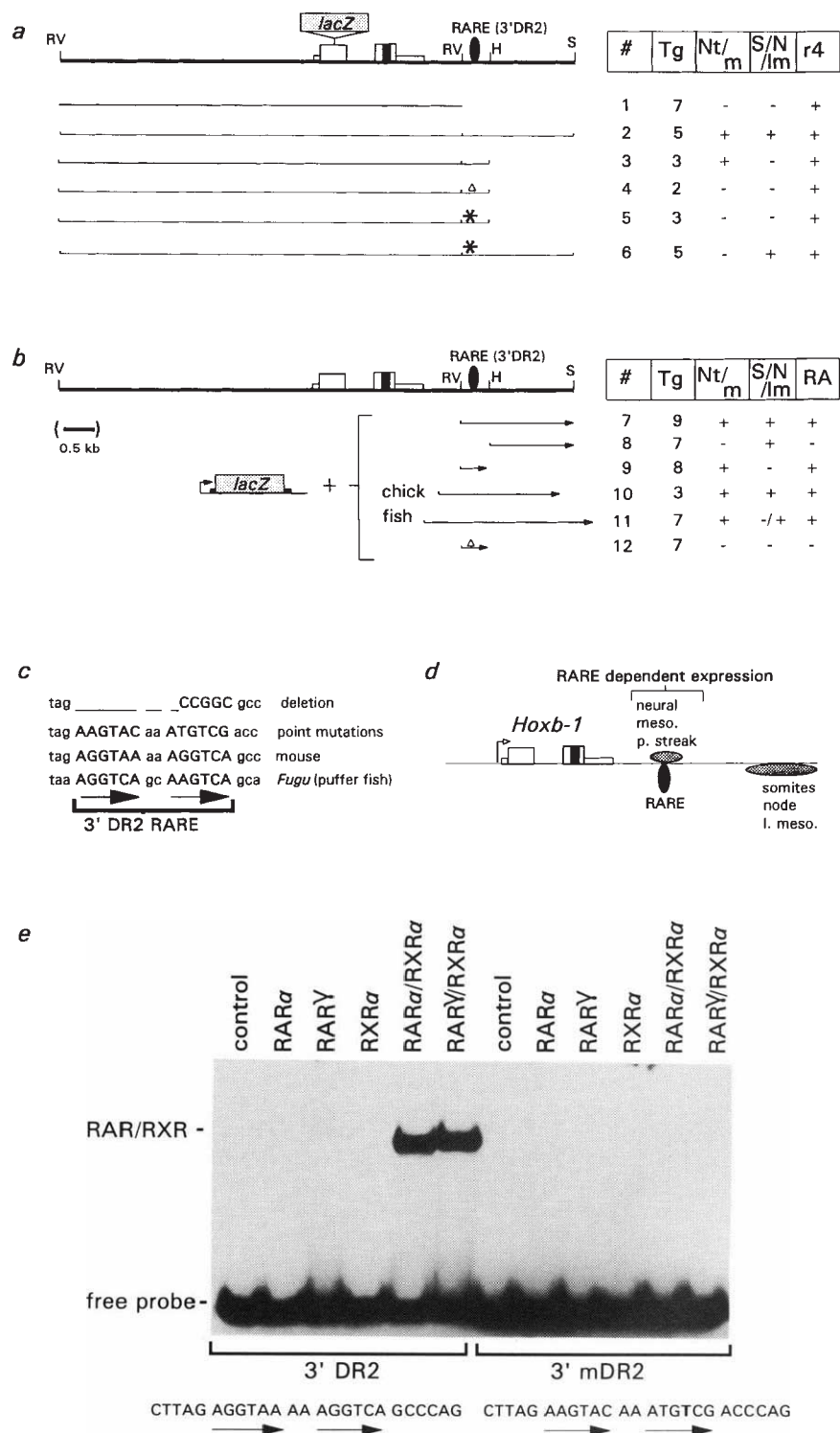
In the absence of the *HindIII-SalI* mesodermal enhancer (construct 3), there was progressive expression in neuroectoderm and lateral mesoderm which extended from the posterior primitive streak into the headfold region near the presumptive rhombomere 3/4 boundary (Fig. 3f-l). This suggests that the 800-base-pair (bp) *EcoRV-HindIII* fragment mediates the remaining domains of early expression of *Hoxb-1* in the ectoderm and a subset of the mesoderm. We tested its activity on a heterologous promoter (construct 9) and found that the region did function as an enhancer, regulating these mesodermal and ectodermal domains (Fig. 3m-g). This enhancer also mediated the early ectodermal/mesodermal response to retinoic acid. Following exposure to retinoic acid at 7.5 d.p.c. there were anterior shifts in the domains of expression (Fig. 3s). Therefore, the two 3' enhancers account for the early endogenous patterns of *Hoxb-1* expression and the response to retinoic acid.

The 800-bp mouse *EcoRV-HindIII* fragment was sequenced to look for potential RAREs and revealed a single motif related to a consensus RARE sequence¹⁷⁻¹⁹, with a spacing of 2 bp between the direct repeats (DR2), while a similar DR2 RARE was located after sequencing the *Fugu* 3' region (Fig. 1c). The ability of the mouse RARE to bind retinoic acid receptors (RARs) and retinoid X receptors (RXRs) was examined in electrophoretic mobility shift assays (Fig. 1e). This sequence bound RAR/RXR heterodimers *in vitro*, but not homodimers, and point mutations in the half sites of the direct repeats abolished binding (Fig. 1c, e).

To investigate the *in vivo* activity of this RARE and to determine whether it is required for the response to retinoic acid in the *EcoRV-HindIII* fragment, we examined a version carrying a deletion (construct 12; Fig. 1c). The deletion abolished early *Hoxb-1* expression at all stages, suggesting that not only is the RARE necessary for the retinoic acid response, but it is also essential in the normal early regulation of the gene. We also tested deletions and point mutations in the RARE, linked to a region that regulates rhombomere 4 expression (constructs 4-6; Fig. 4), and found that they selectively eliminated transgene expression in the lateral mesoderm and ectoderm, leaving only the expression in rhombomere 4 (Fig. 4b, c). In the mutant transgenes there was never any staining in embryos at 7.0-8.25 d.p.c., suggesting that the RARE is required for the establishment, not just the maintenance, of early ectodermal and some mesodermal

§ To whom correspondence should be addressed.

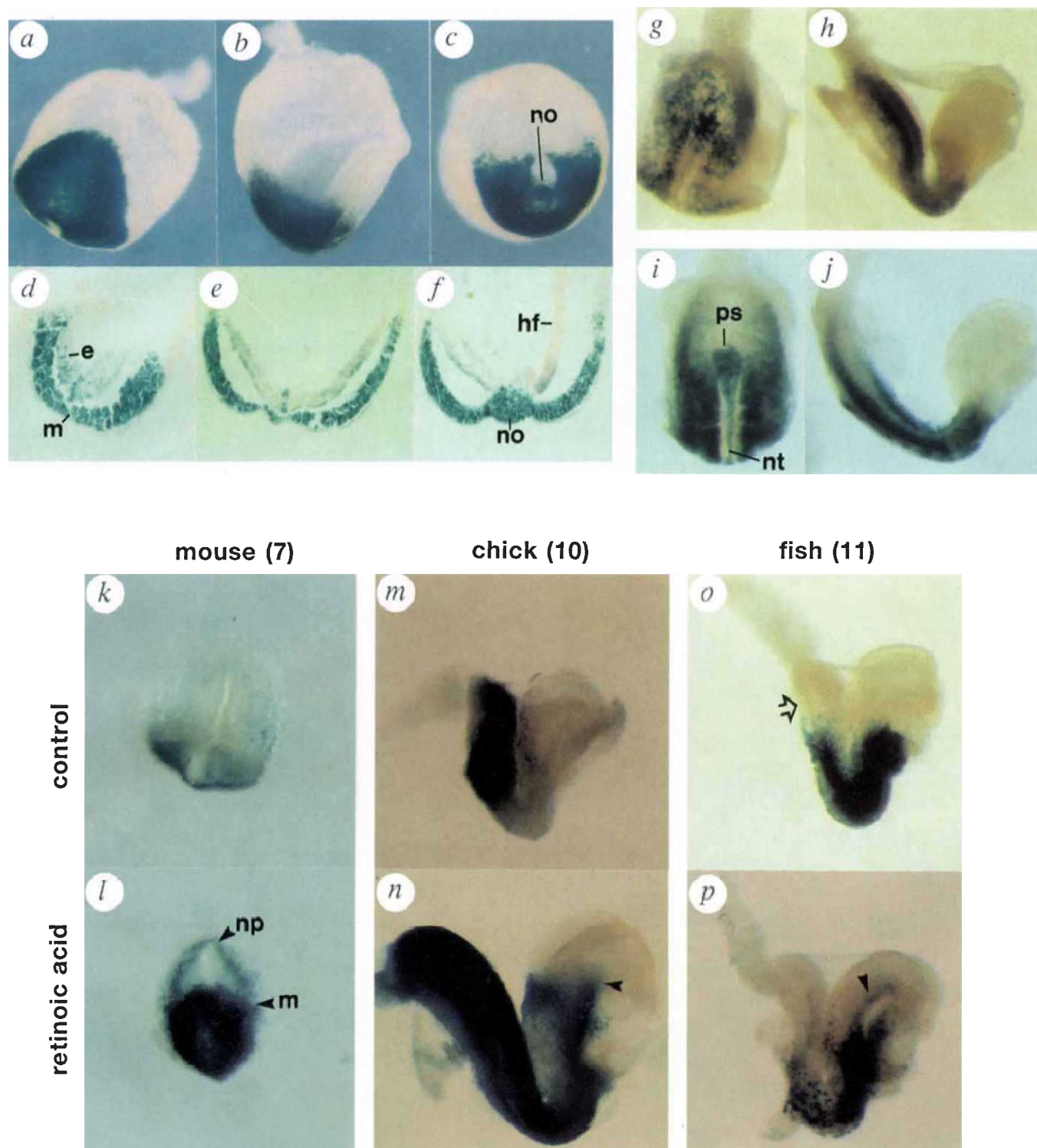
FIG. 1 Characterization of evolutionarily conserved regions that regulate early patterns of *Hoxb-1* expression and its response to retinoic acid. The left side of *a* and *b* shows the *lacZ* reporter constructs used to test for regulatory activity in transgenic mice. The arrows in *b* indicate normal 5' to 3' orientation of *Hoxb-1* fragments on the heterologous promoter. RARE(3' DR2), a retinoic acid response element in which the two direct repeats are separated by 2bp; (Δ), a 9-bp deletion; and *, point mutations in the RARE detailed in *c*. The right side shows the transgenic analysis: #, construct number; Tg, total number of F₀ embryos and lines expressing reproducible patterns for each construct; + or - indicate the presence or absence of transgene expression in neural tube and mesoderm (Nt/m), somites, node and lateral mesoderm, (S/N/lm) rhombomere 4 (r4) and response to retinoic acid (RA). The mouse RARE is located 210 bp from the *EcoRV* site, and in the genes of the mouse and pufferfish, RAREs are positioned in the antisense orientation with respect to the RARE consensus¹⁷⁻¹⁹. *c*, The sequence (sense orientation) of the normal mouse and pufferfish 3' RAREs and deletion or point mutation variants in the mouse RARE used for *in vitro* (e) or transgenic analysis. *d*, Summary of the position and activity of the two enhancers that regulate early *Hoxb-1* expression. The 3' RARE is required for the activity of the enhancer that directs expression in neural tube, mesoderm and primitive streak. *e*, Electrophoretic mobility shift assay (EMSA) *in vitro* binding of RAR and RXR proteins to oligonucleotides spanning the wild-type (3' DR2) and mutant (3' mDR2) forms of the mouse RARE. The receptors added are indicated above each lane and the sequence of the oligonucleotides used are shown below the lanes. There is no binding to the mutant sequence and only RAR/RXR heterodimers can bind to the wild-type RARE. Restriction sites, RV, *EcoRV*; S, *SalI*; H, *HindIII*. **METHODS.** The activity of the *Escherichia coli* β -galactosidase (*lacZ*) reporter gene was assayed at 7.25–9.0 days post-coitum as described previously^{9,28}. Constructs 1–6 contained the *lacZ* gene inserted in-frame in the first exon of the *Hoxb-1* gene at an *Xma*III site in amino acid 34. Constructs 7–10 and 12 used a *lacZ* vector with the human β -globin promoter (pGZ40; ref. 29). Construct 11 had a vector with the *Hoxb-4* promoter linked to *lacZ* (construct 8; ref. 28). The point mutations and deletions in the RARE were generated by inverse polymerase chain reaction followed by complete sequencing of the regions. EMSAs were as described³⁰, using purified bacterially expressed human RAR and RXR proteins provided by P. Chambon.



expression. However, the point mutations in the 3' RARE did not affect the early expression in the somites, notochord and node generated by the *HindIII*–*Sal* fragment (construct 6; Fig. 4e, f), as summarized in Fig. 1d.

As exogenous retinoic acid posteriorizes vertebrate embryos by inducing anterior shifts in gene expression²⁰⁻²⁶, we anticipated that removing target sequences for retinoic acid response would cause a posterior shift in *Hoxb-1* expression. Hence, it is surprising that deletion of the RARE prevented the establishment of neural expression. Together, these results show that a 3' RARE is directly mediating part of the *Hoxb-1* retinoic acid response and is an essential element for establishing a subset of the early

FIG. 2 The early phase of *Hoxb-1* expression is regulated by conserved 3'-flanking sequences containing a DR2 retinoic acid response element. *a–h*, Construct 2 reconstructs the full endogenous *Hoxb-1* expression pattern; posterior (*a*), lateral (*b*) and ventral (*c*) views of whole-mount-stained 7.5 d.p.c. embryo. Expression extends from the most posterior regions to a sharp anterior boundary rostral of the node (no), which also displays staining (*c*). *d–f*, Progressively more anterior serial transverse sections through the embryo shown in *a–c*. There is staining in mesoderm (m) and overlying neural ectoderm (e) in regions posterior to the node (*d, e*). However, mesoderm at this stage has a more anterior limit of expression than ectoderm, because staining extends more rostrally into the headfold (hf) region (*f*). *g, h*, Posterior and lateral views, ►



respectively, of an 8.0-d.p.c. transgenic embryo. Expression at this stage extends to the future rhombomere 3/4 boundary in the neural ectoderm and is regressing posteriorly in the underlying mesoderm. The early pattern of *Hoxb-1* expression is generated by a 3' *EcoRV-SalI* regulatory region as indicated in posterior (i) and lateral (j) views of an 8.0-d.p.c. embryo containing construct 7. There is staining in the notochord (nt) as it emerges from the primitive streak (ps) (g, i). The mouse *Hoxb-1* 3' fragment (construct 7, *EcoRV-SalI*) mediates an anterior shift in neural plate (np) and mesodermal expression within 4 h of retinoic acid treatment at 7.25 d.p.c. (l) compared with control embryos (k). m, o, Conserved 3' regions in the chicken (construct 10, m) and *Fugu* (construct 11, o) *Hoxb-1* genes regulate related early patterns of expression in transgenic mice at 8.0 d.p.c. The open arrow in o denotes lack of

expression in most posterior mesoderm and ectoderm. When exposed to retinoic acid, expression from both the chicken (n) and *Fugu* (p) regulatory domains undergoes an anterior shift. Arrowheads in l, n and p indicate retinoic-acid-induced domains of expression. Numbers in parentheses indicate construct number.

METHODS. Transgenic mice were produced in F₁ backcrossed animals (CBA × C57) and assayed for β -galactosidase activity as described previously^{9,28}. Paraffin sections (6 μ m) were made from whole-mount embryos and stained with eosin²⁸. All-trans-retinoic acid (25 mg ml⁻¹ in dimethyl sulphoxide) was diluted in sesame seed oil and 200 μ l of this was given to pregnant females at 7.25–7.5 d.p.c. by gavage for a final dose of 20 mg kg⁻¹ (ref. 9).

expression patterns during embryogenesis. This requirement for a RARE suggests that endogenous retinoids may be a prerequisite for establishing aspects of *Hoxb-1* expression. We believe that this regulation is an important component of the normal control of the gene, because it has been conserved from *Fugu* to mouse.

Hoxb-1 also has a late response to retinoic acid in the

hindbrain^{8,9,11,12}, and we found that its regulation maps to elements in the 5' region of the gene (M.S. *et al.*, unpublished data). Therefore, the response of *Hoxb-1* to retinoic acid involves multiple control regions and is not dependent on a single master RARE located at the 3' end of the gene. *Hoxa-1*, a *labial* paralogue of *Hoxb-1*, is reported to have a DR5 RARE in the

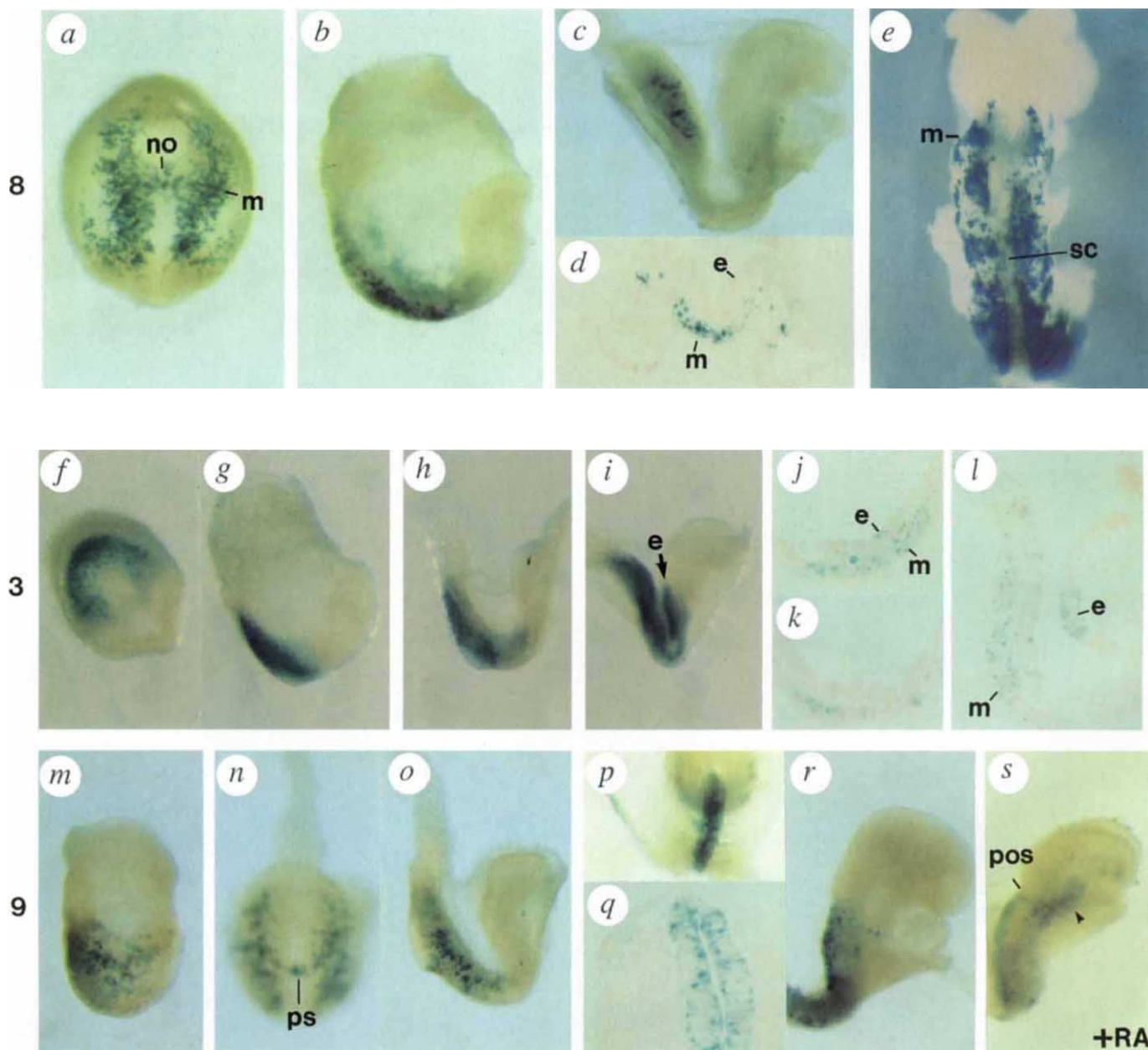
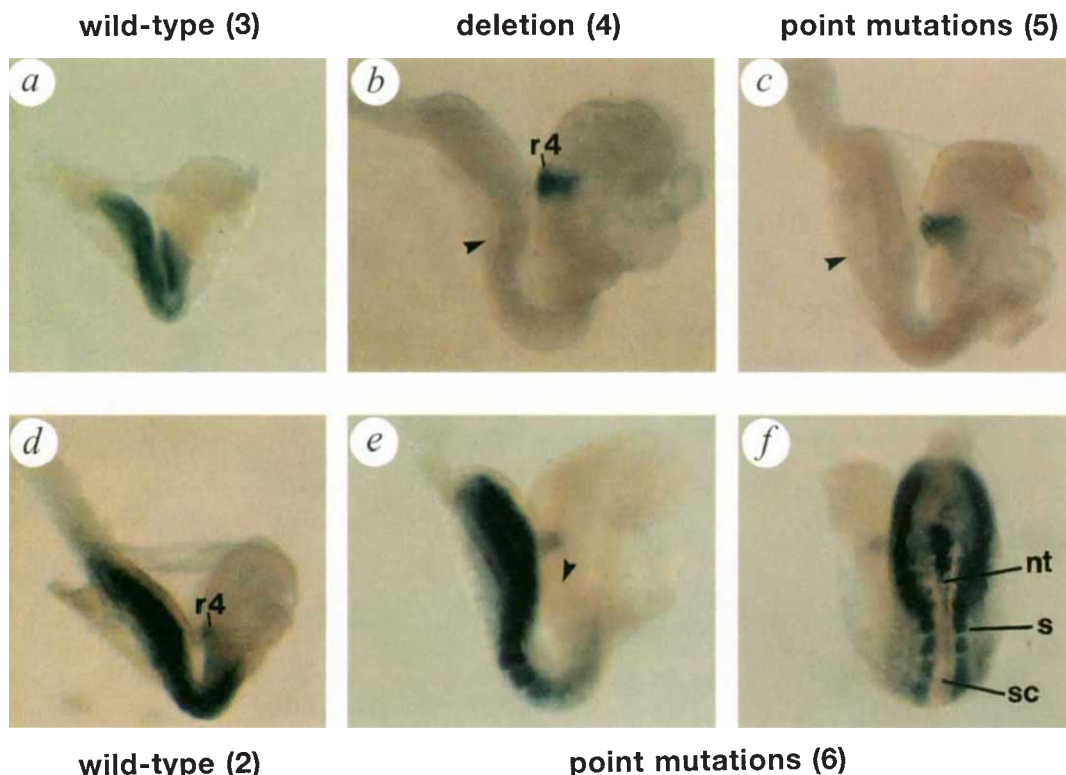


FIG. 3 The mouse 3'-regulatory domain contains two enhancers required for different aspects of early *Hoxb-1* expression. Numbers on the left indicate construct number. *a-e*, Time course of expression using a mesodermal enhancer (construct 8; *HindIII-Sal*). *a, b*, Ventral and lateral views, respectively, of a 7.5-d.p.c. embryo showing that expression is restricted to the node (no), paraxial mesoderm (m) and lateral plate mesoderm. *c, d*, Lateral view and transverse section, respectively, of an 8.5-d.p.c. embryo showing that expression is confined to the mesoderm. *e*, Dorsal view at 9.0 d.p.c. showing staining in somitic mesoderm and not in the spinal cord (sc). *f-s*, Time course of expression using a neural/mesodermal enhancer on the *Hoxb-1* promoter (*f-l*; construct 3) and on a heterologous promoter (*m-s*; construct 9). *f, g*, Posterior and lateral views, respectively, and *j, k*, transverse serial sections of a 7.5 d.p.c. embryo containing construct 3. The most posterior section (*j*) shows staining in both mesoderm (m) and ectoderm (e), although

there is only mesodermal staining in the anterior section (*k*). Lateral views of embryos at 8.0 (*h*) and 8.25 (*i*) d.p.c. show that expression in neural ectoderm now extends more anteriorly to the future rhombomere 3/4 boundary. *l*, Sagittal section of 8.5-d.p.c. embryo in *i* shows mesodermal staining in posterior but not anterior regions, while the ectoderm is stronger in anterior domains. Also shown is expression of construct 9 at 7.0 (*m*), 8.0 (*n, o*) and 8.5 (*r*) d.p.c. Only mesodermal expression is observed at 7.0 d.p.c., and there is a progressive increase in neuroectodermal expression in anterior regions which peaks at 8.5 d.p.c. *p, q*, Dorsal view and transverse section, respectively, of the embryo in *r* showing that anterior expression is confined to the neural ectoderm. The enhancer in the *EcoRV-HindIII* fragment (construct 9) mediates the response to ectopic doses of retinoic acid (RA; *s*), where staining in the mesoderm (see arrowhead) is observed anterior to the preotic sulcus (pos).

FIG. 4 The 3' RARE of the *Hoxb-1* gene is necessary for establishing early neural expression. Numbers in parentheses indicate construct number. a, d, Control staining patterns of transgenic embryos with the wild-type *EcoRV-HindIII* fragment (a) and the *EcoRV-Sal* fragment (d) linked to a construct directing expression in rhombomere 4 (constructs 3 and 2, Fig. 1a). Lateral views of 8.25-d.p.c. embryos containing a deletion (b) or point mutations (c) in the 3' RARE in the *EcoRV-HindIII* enhancer (constructs 4 and 5). At all stages examined (7.25–9.0 d.p.c.) the only expression observed arises in rhombomere 4 after 8.25 d.p.c., indicating that both mutations completely eliminate the early expression in mesoderm and ectoderm. e, f, Lateral and posterior views, respectively, of an embryo carrying point mutations in the 3' RARE in the *EcoRV-Sal* enhancer fragment (construct 6). Posterior expression is present in somitic mesoderm (s) and notochord (nt) but is absent in the spinal cord (sc) and neural tube. Arrowheads indicate domains of expression missing in RARE mutant constructs.



3' region which is necessary for the response to retinoic acid in tissue culture cells²⁷. By analogy to our results with *Hoxb-1*, this RARE could be important for establishing *Hoxa-1* expression, and could reflect a common role for retinoids in regulating vertebrate *labial*-related genes. □

Received 18 April; accepted 24 June 1994.

1. Wilkinson, D. et al. *Nature* **341**, 405–409 (1989).
2. Murphy, P., Davidson, D. & Hill, R. *Nature* **341**, 156–159 (1989).
3. Frohman, M., Boyle, M. & Martin, G. *Development* **110**, 589–607 (1990).
4. Sundin, O. & Eichele, G. *Genes Dev.* **4**, 1267–1276 (1990).
5. Maden, M. et al. *Development* **111**, 35–44 (1991).
6. Guthrie, S. et al. *Nature* **356**, 157–159 (1992).
7. Murphy, P. & Hill, R. *Development* **111**, 61–74 (1991).
8. Conlon, R. & Rossant, J. *Development* **116**, 357–368 (1992).
9. Marshall, H. et al. *Nature* **360**, 737–741 (1992).
10. Sundin, O. & Eichele, G. *Development* **114**, 841–852 (1992).
11. Kessel, M. *Neuron* **10**, 379–393 (1993).
12. Morriss-Kay, G. et al. *EMBO J.* **10**, 2985–2996 (1991).
13. Kessel, M. & Gruss, P. *Cell* **67**, 89–104 (1991).
14. Dolle, P. et al. *Nature* **342**, 767–772 (1989).
15. Izpisua-Belmonte, J.-C. et al. *Nature* **350**, 585–589 (1991).
16. Brenner, S. et al. *Nature* **366**, 265–268 (1993).
17. Leid, M., Kastner, P. & Chambon, P. *Trends biochem. Sci.* **17**, 427–433 (1992).
18. Stunnenberg, H. *Bioessays* **15**, 309–315 (1993).
19. Kiewer, S. A. et al. *Nature* **355**, 446–449 (1992).
20. Durston, A. et al. *Nature* **340**, 140–144 (1989).
21. Ruiz i Altaba, A. & Jessell, T. *Genes Dev.* **5**, 175–187 (1991).
22. Ruiz i Altaba, A. & Jessell, T. *Development* **112**, 945–958 (1991).
23. Papalopulu, N. et al. *Development* **113**, 1145–1159 (1991).
24. Sive, H. & Cheng, P. *Genes Dev.* **5**, 1321–1332 (1991).
25. Sive, H. et al. *Genes Dev.* **4**, 932–942 (1990).
26. Holder, N. & Hill, J. *Development* **113**, 1159–1170 (1991).
27. Langston, A. W. & Gudas, L. J. *Mechanisms of Development* **38**, 217–228 (1992).
28. Whiting, J. et al. *Genes Dev.* **5**, 2048–2059 (1991).
29. Yee, S.-P. & Rigby, P. W. R. *Genes Dev.* **7**, 1277–1289 (1993).
30. Popperl, H. & Featherstone, M. S. *Molec. cell. Biol.* **13**, 257–265 (1993).

ACKNOWLEDGEMENTS. We thank A. Gould for help in generating the point mutations; W. Hatton for help in sectioning embryos; Z. Webster and L. Jones for animal husbandry; P. Chambon and H. Gronemeyer for providing the RAR and RXR proteins for *in vitro* bindings assays; and members of the laboratories for comments and discussion. R.K., H.M. and A.K. were supported by an HFSPO collaborative grant, M.S. by an SNF and EMBO fellowship, H.P. by an EMBO fellowship, and S.A. by an MRC student traineeship.

Crystal structure of *Yersinia* protein tyrosine phosphatase at 2.5 Å and the complex with tungstate

Joanne A. Stuckey*, Heidi L. Schubert*†‡, Eric B. Fauman*, Zhong-Yin Zhang†‡§, Jack E. Dixon†‡ & Mark A. Saper*†¶

* Biophysics Research Division, † Department of Biological Chemistry and ‡ Walther Cancer Institute, The University of Michigan, Ann Arbor, Michigan 48109-1055, USA

PROTEIN tyrosine phosphatases (PTPases) and kinases coregulate the critical levels of phosphorylation necessary for intracellular signalling, cell growth and differentiation^{1,2}. *Yersinia*, the causative bacteria of the bubonic plague and other enteric diseases, secrete an active PTPase³, Yop51, that enters and suppresses host immune cells^{4,5}. Though the catalytic domain is only ~20% identical to human PTP1B⁶, the *Yersinia* PTPase contains all of the invariant residues present in eukaryotic PTPases⁷, including the nucleophilic Cys 403 which forms a phosphocysteine intermediate during catalysis^{3,8–10}. We present here structures of the unliganded (2.5 Å resolution) and tungstate-bound (2.6 Å) crystal forms which reveal that Cys 403 is positioned at the centre of a distinctive phosphate-binding loop. This loop is at the hub of several hydrogen-bond arrays that not only stabilize a bound oxyanion, but may activate Cys 403 as a reactive thiolate. Binding of tungstate triggers a conformational change that traps the oxyanion and swings Asp 356, an important catalytic residue⁷, by ~6 Å into the active

§ Present address: Department of Biochemistry, Albert Einstein College of Medicine, Bronx, New York 10461, USA.

¶ To whom correspondence should be addressed.

site. The same anion-binding loop in PTPases is also found in the enzyme rhodanese¹¹.

The central feature of the *Yersinia* PTPase tertiary fold, as in the structurally homologous PTP1B¹², is a highly twisted, eight-stranded, mixed β -sheet flanked by five α -helices on one side and two α -helices ($\alpha 2$ and $\alpha 3$) on the other (Table 1 and Fig. 1a). At the centre of the β -sheet are two split $\beta\alpha\beta$ motifs¹³ that interdigitate to form four adjacent parallel strands ($\beta 2$, $\beta 8$, $\beta 3$ and $\beta 7$). A prominent depression forming the substrate-binding site is located at the carboxy end of these strands (Fig. 1b). Centred within the site, and located on the β -turn following $\beta 8$ (residues 403–406) and the first turn of helix $\alpha 5$ (407–410), is the PTPase signature sequence (I/V)HCXAGXGR(S/T)(G/A)¹⁴ (single-letter amino-acid code). Residues 403–410 form the PTPase phosphate-binding loop or P-loop with the invariant Cys 403 thiol^{5,9} centred within the loop and close to the $\alpha 5$ helix axis (Figs 1 and 2a).

Although sulphhydryls in proteins normally have a dissociation constant (pK_a) of ~ 8.5 , Cys 403 in *Yersinia* PTPase has an apparent pK_a of 4.7 suggesting that it may be stabilized as a negatively charged thiolate at neutral pH⁹. Active-site thiolates in other hydrolases, such as papain¹⁵ and diene lactone hydrolase¹⁶, are stabilized by ion-pairing to nearby histidines. The only invariant and positively charged residue in the vicinity of Cys 403 is Arg 409, a residue that is important for catalysis¹⁷ (Z.-Y.Z. *et al.*, manuscript in preparation). In the unliganded *Yersinia* PTPase structure, the Arg 409 side chain projects from the $\alpha 5$ amino terminus towards helix $\alpha 3$ and the active-site opening (Figs 1 and 2a). The position of its guanidinium group is anchored by hydrogen bonds to the Leu 285 carbonyl and Thr 358 O γ 1, and by a bidentate salt bridge to the carboxylate of Glu 290, the putative catalytic base⁷ (Fig. 2c). Cys 403 is clearly in the g^+ ($\chi_1 \approx +60^\circ$) rotamer (Fig. 2a), and S γ is 5–6 Å from the Arg 409 guanidinium group. In the unliganded structure of PTP1B solved at 2.9 Å resolution¹², the equivalent

Cys 215 S γ is in a different rotamer conformation that places S γ within 3.0 Å of the Arg 221 guanidinium group. This was proposed to partially stabilize the negatively charged thiolate¹². On the basis of our electron density maps at 2.5 Å resolution (Fig. 2a), we conclude that Arg 409 does not play a major role in stabilizing the thiolate in *Yersinia* PTPase.

In *Yersinia*, the thiol anion of Cys 403 appears to be stabilized by an extensive network of hydrogen bonds¹⁸ (Fig. 2c). The closest proton donor is the O γ 1 hydroxyl of conserved Thr 410 (Thr or Ser in all PTPases). The corresponding Ser 222 in the PTP1B structure¹² does not make this interaction, but hydrogen bonds to the amide of Gly 218. This is probably due to the different cysteine orientation in PTP1B mentioned above.

Cys 403 S γ is also located between 3.0 and 4.2 Å from every amide nitrogen of the P-loop (residues 403–410). At least five of these are shorter than 4.0 Å and should make reasonable S–HN hydrogen bonds¹⁹. Alternatively, the P-loop amides may be considered individual microdipoles with their δ^+ ends²⁰ all oriented towards the thiolate. The P-loop peptide nitrogens are coupled electrostatically through their carbonyl oxygens¹⁸ to several hydrogen-bonding arrays radiating out from the P-loop (Fig. 2c). The invariant (and mostly buried) residues His 402, Tyr 301, Arg 437, Arg 440 and Asn 245 participate in three of these arrays. The first array begins with the Cys 403 carbonyl, and continues through the His 402 imidazole ring, the Tyr 301 hydroxyl, and finally the His 270 side chain (Fig. 2c). Site-directed mutations of the invariant His 402 residue to Asn or Ala raise the apparent pK_a of Cys 403⁹, probably by partially disrupting this hydrogen-bond network. In the second array, peptide bonds following residues 405 and 406 mediate ionic interactions between the thiolate and the buried guanidinium group of Arg 440. Mutation of Arg 440 to Ala or Lys severely impairs catalytic efficiency and binding (Z.-Y.Z. *et al.*, manuscript in preparation). The third array couples Val 407 O to the Arg 437 side chain, but is mediated through an additional

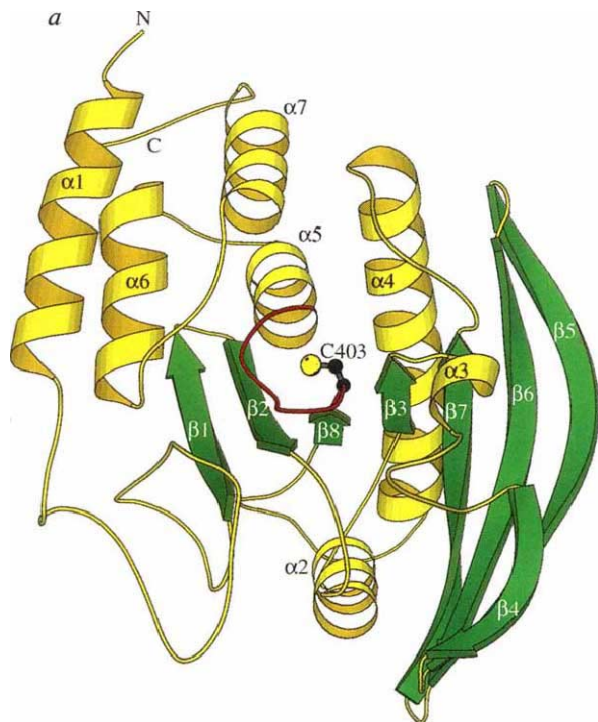
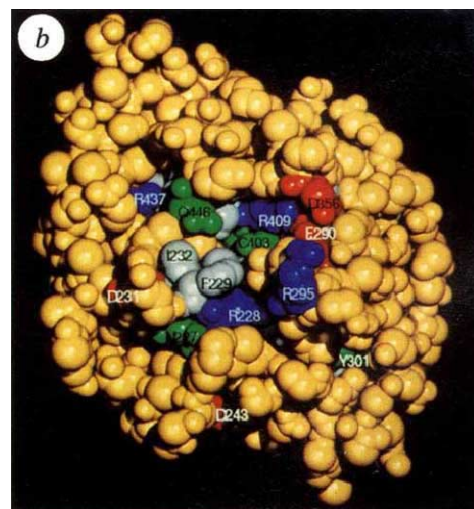


FIG. 1 a, Ribbon diagram of the unliganded *Yersinia* PTPase structure (drawn with MOLSCRIPT³⁰). Molecule is oriented with the C-terminal end of the $\beta 8$ strand towards the viewer (α -helices and turns in yellow, β -strands in green). The catalytic site is centred around the anion-binding loop (P-loop) formed between $\beta 8$ and $\alpha 5$ (residues 403–410, red; Cys 403, yellow). Surrounding the binding site are residues contributed by turns $\alpha 1$ – $\beta 1$, $\alpha 6$ – $\alpha 7$, $\beta 2$ – $\alpha 2$, $\alpha 3$ – $\beta 4$ and $\alpha 4$ – $\beta 7$, and helix



$\alpha 3$. b, CPK model of unliganded PTPase in the same orientation as a, highlighting the conserved and invariant residues⁷ that are solvent accessible. Polar uncharged residues, green; basic, blue; acidic red; and hydrophobic, grey. Positions of unlabelled conserved residues: Gly 408, grey, upper-left of Cys 403; Gln 450, green, behind Gly 408 and Arg 409; Val 433, Met 444, Val 445, grey, in pocket adjacent to Arg 437. Only the main-chain of Tyr 301 is accessible; the side chain is buried and interacts with His 402 and His 270 in the hydrogen-bonding network (Fig. 2c). As noted for the PTP1B structure¹², most of the residues clustered within or around the catalytic depression are invariant of structurally similar within the PTPase family⁷. This suggests that all PTPases are likely to recognize phosphotyrosine in a similar way and have similar catalytic mechanisms.

peptide bond (444–445). Because mutations of the invariant Arg 437 have little effect on catalysis or oxyanion binding (Z.-Y.Z. *et al.*, unpublished data), we suggest that this residue might interact with conserved functional groups (such as main-chain or anionic side-chain atoms) of a phosphorylated peptide

substrate. The Arg 437 guanidinium group is located ~ 14 Å from the active site, but is partially exposed in a predominantly hydrophobic pocket between $\alpha 1$ and $\alpha 7$ (Fig. 1b). Finally, the amides of P-loop residues 408 and 409 on the first turn of the $\alpha 5$ helix are coupled by hydrogen bonds to peptide bonds on

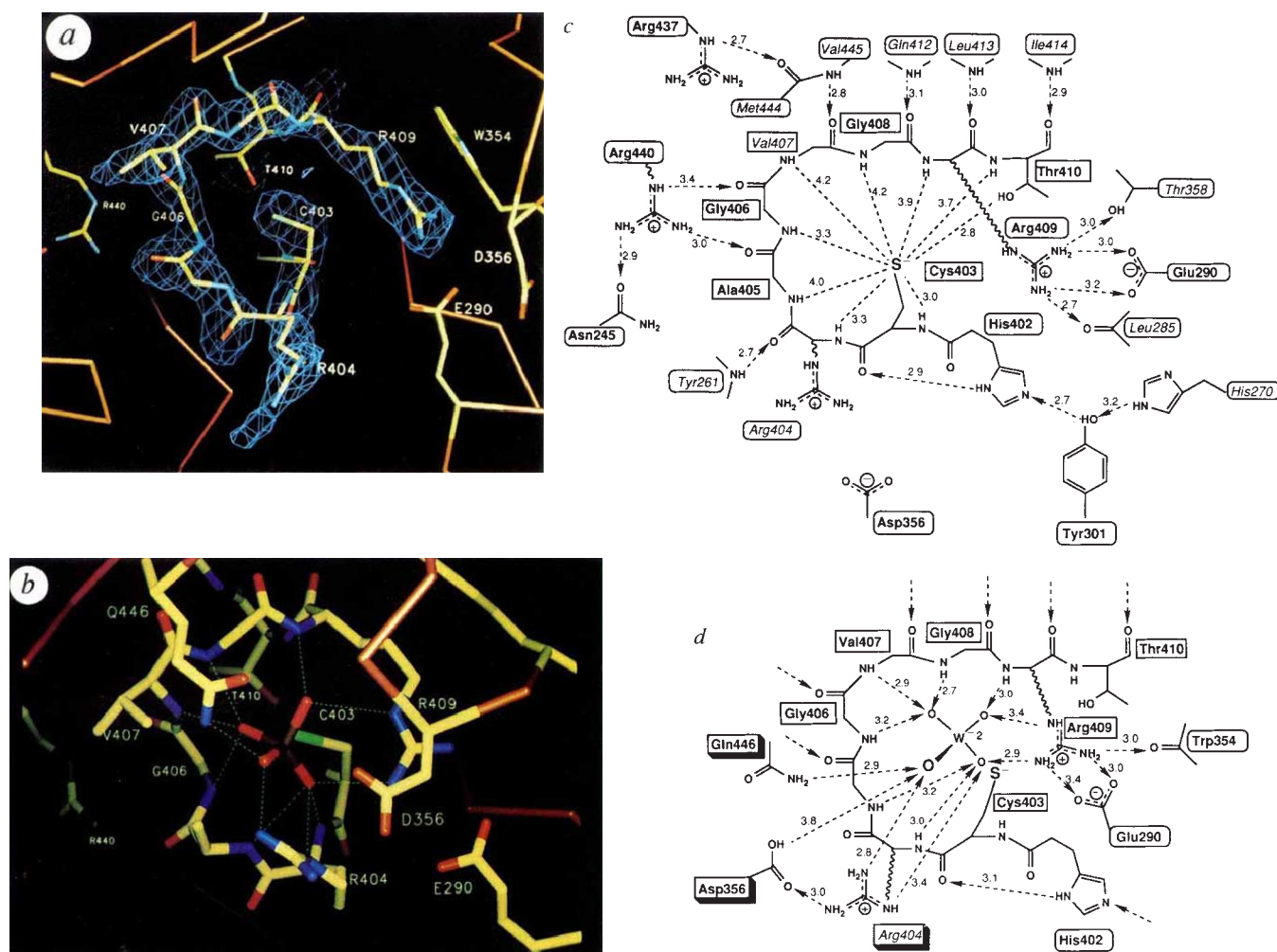


FIG. 2 The *Yersinia* PTPase P-loop (residues 403–410) and hydrogen-bonding arrays. *a* and *b*, Carbons, nitrogens, oxygens, sulphurs and the tungsten are shown in yellow, blue, red, green and violet, respectively. *a*, Unliganded *Yersinia* PTPase structure. Superimposed on the P-loop are electron density contours (2.5 σ) from an 'omit' map calculated with coefficients ($F_{\text{obs}} - F_{\text{calc}}$), α_{calc} . *b*, PTPase-tungstate structure. Hydrogen bonds (3.5 Å or less) between the tungstate oxygens and the enzyme are shown as dashed green lines. As described in the text, the $\beta 7$ – $\alpha 4$ loop containing Asp 356 has moved to interact with the bound oxyanion. The corresponding $\beta 11$ – $\alpha 3$ loop in both unliganded and tungstate-bound structures of PTP1B¹² appears to be in a position similar to the *Yersinia* PTPase unliganded state with the analogous Asp 181 quite distant from the bound tungstate. Crystal lattice restrictions in PTP1B may have inhibited a conformational change because the PTP1B-tungstate complex was prepared by soaking unliganded crystals¹², whereas the *Yersinia* PTPase- WO_4^{2-} crystals were co-crystallized with tungstate. Alternatively, sequence differences in PTP1B may make the loop movement energetically less favourable, and account for the enzyme's lower catalytic rate when compared to *Yersinia* PTPase³¹. *c*, *d*, Active-site hydrogen-bonding arrays in the *Yersinia* PTPase structures. Arrows point from hydrogen donor to acceptor. Residues involved in primary hydrogen bonds with the Cys 403 thiolate or bound tungstate are boxed. Conserved and invariant residues are shown in bold and non-conserved residues in italics. All distances in Å. *c*, Cys 403 is stabilized as an anion by several hydrogen-bonding arrays radiating from the P-loop (see text). Sulphur–nitrogen interactions under 4.0 Å are expected to form reasonable S–HN hydrogen bonds¹⁹. Cys 403 S γ is ~ 10 Å from Asp 356, and 6 Å from Arg 409 N ϵ . *d*, PTPase- WO_4^{2-}

structure showing hydrogen bonds to the tungstate oxygens. All secondary hydrogen-bonding arrays shown in *c* also exist in *d*, but have been omitted for clarity. The oxygen of WO_4^{2-} indicated with an enlarged O is expected to correspond to the oxygen of the scissile O–P bond in a phosphorytyrosine substrate. Glu 290 O $\epsilon 2$ is 5.5 Å from the nearest tungstate oxygen; Cys 403 S γ is 3.6 Å from the tungsten. METHODS. Calculated phases for the difference map in *a* were obtained from a simulated-annealing X-PLOR refinement with residues 403–410 omitted from the structure factor calculations. *b* and *d*, Crystals of the PTPase-tungstate complex ($\text{P}_2\text{I}_2\text{I}_1$, $a = 56.4$ Å, $b = 49.8$ Å, $c = 100.8$ Å; size $0.8 \times 0.2 \times 0.02$ mm³) were grown by mixing equal volumes of a PTPase solution containing 1 mM tungstate, with precipitant (22% (w/v) polyethylene glycol 4000, 200 mM Li_2SO_4 , 10% isopropanol, 0.1% β -mercaptoethanol, 100 mM Tris-HCl, pH 8.5). The crystal diffracted to at least 2.2 Å, and 81% of the reflections between 7.0 Å and 2.6 Å were recorded. These amplitudes (and those from a Hg derivative) were used to phase data from isomorphous crystals of the *Yersinia* PTPase Cys 403 $\rightarrow$ Ser mutant containing bound sulphate (H.L.S. *et al.*, manuscript in preparation). After most of the Cys 403 $\rightarrow$ Ser structure had been built, a tetrahedral tungstate ion (net -2 charge, W–O bond length = 1.77 Å) was fit into a difference map, and the entire model further refined with X-PLOR²⁹ against the PTPase-tungstate data. Force-field interactions between the tungstate and Cys 403 were turned off to allow close-approach of these two groups if necessary. The PTPase- WO_4^{2-} structure currently has a crystallographic R -factor of 19.5% to 2.6 Å ($F/\sigma_F \geq 2.0$). The Ca r.m.s. deviation from the unliganded structure is 0.9 Å. A detailed description of the structure refined to 2.5 Å will be published elsewhere (E.B.F. *et al.*, manuscript in preparation).

TABLE 1 Diffraction data, phasing, density averaging and refinement statistics for unliganded *Yersinia* PTPase

Compound*	Resolution (Å)	No of unique reflections	Completeness (%)	$R_{\text{merge}}^{\dagger}$ (%)	$R_{\text{iso}}^{\ddagger}$ (%)	No. of sites	Phasing power§ (max. resolution used in phasing, Å)
Native	2.5	19,476	91	5.9			NA
$\text{K}_2\text{Pt}(\text{CN})_4$	2.3	19,501	72	9.6	15.7	1	1.42 (2.8)
$\text{K}_2\text{Pt}(\text{CN})_4$	2.3	12,793	46	10.0	13.9	1	1.52 (2.8)
PtCl_4	2.8	11,106	75	8.0	16.2	1	1.26 (3.5)
K_2PtCl_4	3.0	9,133	77	10.6	20.2	2	1.38 (3.5)
Overall mean figure of merit (20–2.8 Å) = 0.51							

The catalytic domain (residues 163–468) of Yop51, a PTPase from *Yersinia enterocolitica*, was expressed, purified, and crystallized as described previously²⁴. Trigonal crystals (space group $P3_2$; $a = b = 71.53$ Å, $c = 107.48$ Å, $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$; two molecules per asymmetric unit) grew only in the absence of oxanion-containing salts. All diffraction data were collected on a SDMS multiwire area detector system mounted on a Rigaku RU-200 rotating anode generator (50 kV, 100 mA), and processed by the accompanying software²⁵. Single isomorphous replacement phases, derived from the first $\text{K}_2\text{Pt}(\text{CN})_4$ derivative, were improved by solvent flattening using PHASES²⁶. Sites for the other heavy-atom derivatives were located in difference Fourier maps. Multiple isomorphous replacement (MIR) phases (20–3.5 Å resolution; mean, figure of merit, 0.56) were improved by iterative solvent flattening, and an initial backbone trace of one molecule was fitted to the resulting density map with the program O²⁷. The location of the non-crystallographic 2-fold axis was found by a rotation/translation search of the MIR map (PROTEIN²⁸), and refined by PHASES (density correlation value, 0.32). Over 60% of the model (mainly polyalanine) was fit to a 20–3.0 Å MIR map (mean figure of merit = 0.51) that had been averaged without a mask. Masks derived from this model were used to improve the MIR map by iterative density averaging and solvent flattening²⁶. PTPase models were refined with the 'slow-cool', simulated annealing (3,000 K) protocol of X-PLOR²⁹ (including non-crystallographic restraints), and refit to 3.0 Å electron density maps calculated with MIR-combined phase probabilities, and later at 2.8 Å and 2.5 Å with $(2F_{\text{obs}} - F_{\text{calc}})$, α_{calc} maps. The R -factor is currently 18.4% (free R -factor²⁹ = 27.0%) for 17,095 unique reflections in the 7–2.5 Å shell ($F/\sigma_F \geq 2$, 84% complete). The model includes residues 189–468 and 191–468 for molecules A and B, respectively, a total of 4,290 non-hydrogen atoms. There was no interpretable electron density for the N-terminal residues 163–188 in either molecule and they are assumed to be disordered. The r.m.s. deviations for bond distances and angles were 0.015 Å and 3.3° , respectively. Refinement to higher resolution, relaxation of the non-crystallographic restraints, and inclusion of solvent molecules is in progress. Coordinates will be submitted to the Brookhaven Protein Data Bank.

* Derivative crystals were soaked in solutions containing 0.05–0.1 mM heavy-atom reagent in 22% polyethylene glycol 1500, 10% 2-methyl-2,4-pentanediol, 1 mM imidazole, pH 7.2, for up to 20 h before data collection.

$\dagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i - \langle I \rangle| / \sum \langle I \rangle$ where I_i is the intensity measurement of a particular symmetry-related reflection.

$\ddagger R_{\text{iso}} = \sum (|F_{\text{PH}}| - |F_P|) / \sum |F_P|$, where F_P and F_{PH} are the structure factor amplitudes of the native protein and protein complexed with heavy atoms, respectively.

§ Phasing power is defined as r.m.s. $F_h / (\text{r.m.s. lack of closure error})$, where F_h is the heavy-atom structure factor.

|| Data from two $\text{K}_2\text{Pt}(\text{CN})_4$ soaked crystals were treated as separate heavy-atom derivatives for multiple isomorphous replacement phasing.

the second turn of the helix. Though previously attributed to a helix macrodipole effect²¹ in the PTP1B structure¹², the polarized microdipoles of peptide bonds on the first two turns of an α -helix (such as in the P-loop) have been shown to confer the majority of the charge stabilization energy²⁰.

To examine how phosphoryl groups of substrates would bind to PTPases, we transferred the unliganded PTPase crystals to solutions containing 0.01–1.0 mM sodium tungstate, a competitive inhibitor (dissociation constant $K_D = 61 \mu\text{M}$ at pH 7; Z.-Y. Z. and J. E. D., unpublished data). The crystals cracked

within hours. By adding 1 mM tungstate to the enzyme before crystallization, we obtained a different crystal form containing the PTPase- WO_4^{2-} complex and have independently solved its structure to 2.6 Å resolution (Fig. 2b; E.B.F. *et al.*, manuscript in preparation). The P-loop orients the bound tungstate directly over the nucleophilic Cys 403 ($\text{S}\gamma$ -W distance = 3.6 Å) by anchoring three of the anion's oxygens to the main-chain amides of residues 404–409 through hydrogen bonds (Fig. 2d; mean distance, 3.0 Å). The -2 charge on the tungstate, or a -3 charge of the pentacoordinated transition state expected during

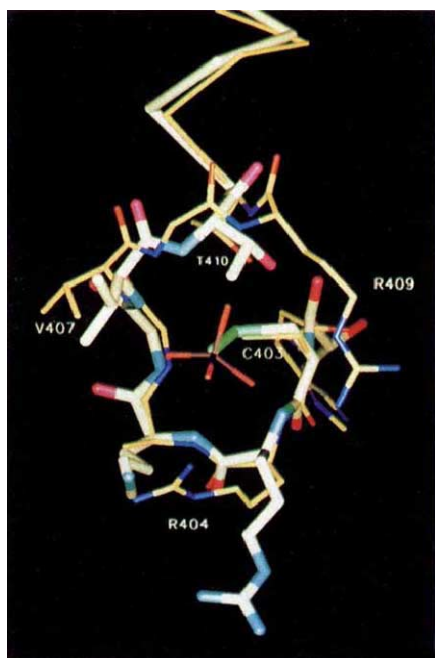


FIG. 3 Superposition of the *Yersinia* PTPase P-loop structure onto the anion-binding loop of rhodanese. The PTPase-tungstate complex is in yellow with thin bonds, and rhodanese is white with thick bonds. Residue numbers are for the PTPase structure. The corresponding sequence alignment for the anion-binding loops is:

PTPase- WO_4^{2-} 403 **C R A G V G R T A** 411

Rhodanese 247 **C R K G V T . A C** 254

Arg 409, catalytically important in all PTPases, is an insertion relative to rhodanese, which in effect widens the first turn of $\alpha 5$ in the PTPase. In both structures, the main-chain NH groups from 403–409 are all oriented towards the centre of the anion-binding loop. Though functionally unrelated to PTPases, the catalytic mechanism is similar. Rhodanese transfers a sulphur from thiosulphate ($\text{S}_2\text{O}_3^{2-}$) to a nucleophilic thiolate to form a charged, persulphide ($\text{Cys 247-S}\gamma\text{-S}\delta^-$) enzyme intermediate as depicted here. When superimposed on the PTPase P-loop, the covalently-bound sulphur anion ($\text{S}\delta^-$) is just 1.6 Å from the bound tungsten. In rhodanese, cyanide (or possibly a non-haem iron protein) binds and accepts S to form thiocyanate (SCN^-)¹¹. In PTPases, a hydroxyl ion is the acceptor forming HPO_4^{2-} .

METHODS. Rhodanese coordinates¹¹ (PDB code 1RHD) were superimposed on the PTPase- WO_4^{2-} structure by minimizing the $\text{C}\alpha$ r.m.s. deviation between residues 398–407, 410–421 and 254–260 in PTPase- WO_4^{2-} , and 242–251, 253–264 and 266–272 in rhodanese. The r.m.s. deviation for these 29 pairs of α -carbons was 1.4 Å.

phosphate hydrolysis, will be delocalized by the same hydrogen-bond arrays discussed above in the unliganded structure. The invariant Arg 409 guanidinium group moves ~ 2 Å to form a bifurcated salt bridge with the tungstate oxygens while maintaining its salt bridge with Glu 290 (Fig. 2d). Because mutation of Arg 409 to Lys severely impairs catalysis, we suspect that this bifurcated interaction is essential to bind preferentially the transition state that occurs during phosphate hydrolysis (Z.-Y.Z. *et al.*, manuscript in preparation).

The major difference between the PTPase- WO_4^{2-} complex and the unliganded structure is that the $\beta 7$ - $\alpha 4$ loop (residues 350–360) has moved an average of 3.3 Å and folds over the active site to sequester the bound oxyanion (Fig. 2a, b; H.L.S. *et al.*, manuscript in preparation). In the unliganded PTPase structure, the putative catalytic general acid⁷ Asp 356 is ~ 10 Å from the phosphate-binding site, and interacts with Ser 287 and Ser 289. In the tungstate complex, Asp 356 *Ca* moves towards the active site by 6 Å, hydrogen-bonds with Arg 404, and positions its carboxylate 3.5–3.8 Å away from the tungstate oxygen that is not ligated to the P-loop (large O of tungstate in Fig. 2d). Because this oxygen is directly opposite the Cys 403 nucleophile it probably corresponds to the atom of the scissile ester bond in a phosphotyrosine residue. Assuming this conformational change also occurs on binding of a phosphotyrosyl substrate, Asp 356 would be ideally positioned to aid proton transfer during one or both of the proposed hydrolysis steps¹⁴. The 2.5 Å structure of the inactive Cys 403→Ser PTPase mutant has a sulphate anion bound to the P-loop and displays the identical conformational change (H.L.S. *et al.*, manuscript in preparation). The reported structures of tungstate-free and tungstate-bound forms of the PTP1B structure¹² do not exhibit the above conformational change (discussed in Fig. 2).

Unexpectedly, we discovered a striking similarity between the active-site structures of *Yersinia* PTPase and that of rhodanese, a sulphur transferase found in a variety of organisms^{11,22,23}. Not only are both anion-binding loop conformations very similar (Fig. 3) but they also have an identically positioned cysteine thiolate which is stabilized by hydrogen bonds to amides of the P-loop, and by the amino-end of an α -helix¹¹. Each enzyme also forms a charged covalent intermediate during catalysis, Cys 403– $\text{S}\gamma$ - PO_3^{2-} in PTPase and Cys 247– $\text{S}\gamma$ - $\text{S}\delta^-$ in rhodanese, with the anion positioned at the centre of the P-loop. Additionally, rhodanese Arg 110 (conserved in all rhodanese sequences) is buried and forms hydrogen bonds to the same P-loop amides as the buried Arg 440 in PTPases. We might expect other unrelated enzyme structures to also use the P-loop motif for anion binding and transfer. □

Received 5 April; accepted 28 June 1994.

1. Fischer, E. H., Charbonneau, H. & Tonks, N. K. *Science* **253**, 401–406 (1991).
2. Walton, K. M. & Dixon, J. E. *Rev. Biochem.* **62**, 101–120 (1993).
3. Guan, K. L. & Dixon, J. E. *Science* **249**, 553–556 (1990).
4. Bolin, I. & Wolf-Watz, H. *Molec. Microbiol.* **2**, 237–245 (1988).
5. Bliska, J. B., Guan, K. L., Dixon, J. E. & Falkow, S. *Proc. natn. Acad. Sci. U.S.A.* **88**, 1187–1191 (1991).
6. Tonks, N. K., Diltz, C. D. & Fischer, E. H. *J. biol. Chem.* **263**, 6722–6730 (1988).
7. Zhang, Z. Y., Wang, Y. & Dixon, J. E. *Proc. natn. Acad. Sci. U.S.A.* **91**, 1624–1627 (1994).
8. Guan, K. L. & Dixon, J. E. *J. biol. Chem.* **266**, 17026–17030 (1991).
9. Zhang, Z. Y. & Dixon, J. E. *Biochemistry* **32**, 9340–9345 (1993).
10. Cho, H. *et al.* *J. Am. chem. Soc.* **114**, 7296–7298 (1992).
11. Ploegman, J. H. *et al.* *Nature* **273**, 124–129 (1978).
12. Barford, D., Flint, A. J. & Tonks, N. K. *Science* **263**, 1397–1404 (1994).
13. Orengo, C. A. & Thornton, J. M. *Structure* **1**, 105–120 (1993).
14. Zhang, Z. Y. & Dixon, J. E. *Adv. Enzym. Relat. Areas molec. Biol.* **66**, 1–36 (1994).
15. Lewis, S. D., Johnson, F. A. & Shafer, J. A. *Biochemistry* **20**, 48–51 (1981).
16. Pathak, D. & Ollis, D. J. *molec. Biol.* **214**, 497–525 (1990).
17. Strulli, M., Krueger, N. X., Thai, T., Tang, M. & Saito, H. *EMBO J.* **9**, 2399–2407 (1990).
18. Quiocho, F. A., Sack, J. S. & Vyas, N. K. *Nature* **329**, 561–564 (1987).
19. Gregoret, L. M., Rader, S. D., Fletterick, R. J. & Cohen, F. E. *Proteins* **9**, 99–107 (1991).
20. Åqvist, J., Luecke, H., Quiocho, F. A. & Warshel, A. *Proc. natn. Acad. Sci. U.S.A.* **88**, 2026–2030 (1991).
21. Hol, W. G., van Duijnen, P. T. & Berendsen, H. J. *Nature* **273**, 443–446 (1978).
22. Ploegman, J. H., Drent, G., Kalk, K. H. & Hol, W. G. *J. molec. Biol.* **127**, 149–162 (1979).
23. Wierenga, R. K. *Trans. Am. Crystallogr. Ass.* **22**, 49–62 (1986).
24. Zhang, Z. Y. *et al.* *J. biol. Chem.* **267**, 23759–23766 (1992).
25. Howard, A. J., Nielsen, C. & Xuong, N. H. *Meth. Enzym.* **114**, 452–472 (1985).
26. Furey, W. & Swaminathan, S. *Am. Crystallogr. Ass. Meeting Abstr.* **18**, 73 (1990).

27. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. *Acta crystallogr.* **A47**, 110–119 (1991).
28. Steigemann, W. thesis, Technische Universität, Munich (1974).
29. Brünger, T. A. *XPLOR Version 3.1 Manual* (Yale Univ. Press, New Haven, 1993).
30. Kraulis, P. J. *J. appl. Crystallogr.* **24**, 946–950 (1991).
31. Zhang, Z. Y., Maclean, D., McNamara, D. J., Sawyer, T. K. & Dixon, J. E. *Biochemistry* **33**, 2285–2290 (1994).

ACKNOWLEDGEMENTS. E.B.F. and H.L.S. contributed equally to these results. We thank members of M. Ludwig's and J. Dixon's laboratories, and K. Shaw for their expertise; D. Hume, J. Clement and J. Clemens for their contributions; F. E. Cohen and J. F. Bazan for discussions; and W. Furey, J. Pflugrath, W. Steigemann, T. A. Jones and P. Kraulis for providing software. Financial support provided by the American Cancer Society (J.A.S.), U-M Multipurpose Arthritis Center (M.A.S.), U-M Program in Protein Structure and Design (M.A.S.), Parke-Davis/Warner Lambert (M.A.S.), the NIH (J.E.D., M.A.S.) and Walther Cancer Institute (J.E.D.). H.L.S. is a Joe Dawson Fellow. M.A.S. is a Pew Scholar in the Biomedical Sciences.

The crystal structure of a low-molecular-weight phosphotyrosine protein phosphatase

Xiao-Dong Su*, Niccolo' Taddei†, Massimo Stefani†, Giampietro Ramponi† & Pär Nordlund*

* Department of Molecular Biology, University of Stockholm, Stockholm, Sweden

† Department of Biochemical Sciences, University of Florence, Florence, Italy

PROTEIN tyrosine phosphorylation and dephosphorylation are central reactions for control of cellular division, differentiation and development¹. Here we describe the crystal structure of a low-molecular-weight phosphotyrosine protein phosphatase (PTPase)², a cytosolic phosphatase present in many mammalian cells. The enzyme catalyses the dephosphorylation of phosphotyrosine-containing substrates^{3–6}, and overexpression of the protein in normal and transformed cells inhibits cell proliferation^{7,8}. The structure of the low-molecular-weight PTPase reveals an α/β protein containing a phosphate-binding loop motif at the amino end of helix $\alpha 1$. This motif includes the essential active-site residues Cys 12 and Arg 18 and bears striking similarities to the active-site motif recently described in the structure of human PTP1B⁹. The structure of the low-molecular-weight PTPase supports a reaction mechanism involving the conserved Cys 12 as an attacking nucleophile in an in-line associative mechanism. The structure also suggests a catalytic role for Asp 129 in the reaction cycle.

The bovine liver low-molecular-weight PTPase contains 157 amino-acid residues (M_r 17,953)²; it has been crystallized in 0.1 M acetate buffer, pH 5.5, in the presence of 40% saturated ammonium sulphate¹⁰. The structure of the enzyme was solved using the multiple isomorphous replacement method combined with density modification techniques (Table 1). A model containing residues 4 to 157, one sulphate ion and 83 water molecules has been built and subsequently refined at 2.1 Å resolution. The present model has a crystallographic *R*-factor of 15.6% and shows good stereochemistry (Table 1).

The structure of the protein is composed of a single α/β domain containing a central four-stranded, parallel open-twisted β -sheet surrounded by connecting α -helices on both sides (Fig. 1a). The secondary structure topology is shown in Fig. 1b. Residues 13–17 form a smooth loop connecting strand $\beta 1$ and helix $\alpha 1$. A strong density at the amino end of helix $\alpha 1$ has been interpreted as a sulphate ion (Fig. 2).

In the refined structure, this sulphate ion forms an extensive hydrogen-bonding network to the amide peptide groups of the

loop, as well as to the side chain of Cys 12 and an electrostatic bond to the guanidinium group of Arg 18 (Fig. 3). Cys 12 and Arg 18 have been shown by oligonucleotide-directed mutagenesis to be essential catalytic residues¹¹. Furthermore, one oxygen of the sulphate ion forms a hydrogen bond with Asp 129. This hydrogen bond requires a protonation of either Asp 129 or of the sulphate ion. Asp 129 is probably protonated because the negative charge of the sulphate ion could shift the pK_a of Asp 129 from 3.9 in the free amino acid to over 5.5, which is the pH of the crystals.

The sulphate-binding loop of the PTPase without doubt corresponds to the binding site of the phosphate group of the phosphotyrosine-containing substrates. The sulphate-binding site is found at the bottom of a crevice formed on the C-terminal side of the parallel β -sheet. The edge of this cavity is lined mainly with hydrophobic residues: Leu 13, Ile 16, Trp 49, Asn 50, Tyr 131 and Tyr 132 (Fig. 3). These residues may form a suitable interaction surface for the aromatic moiety of the phosphotyrosine-containing substrate, as well as a binding site for neighbouring amino acids in the peptide substrate.

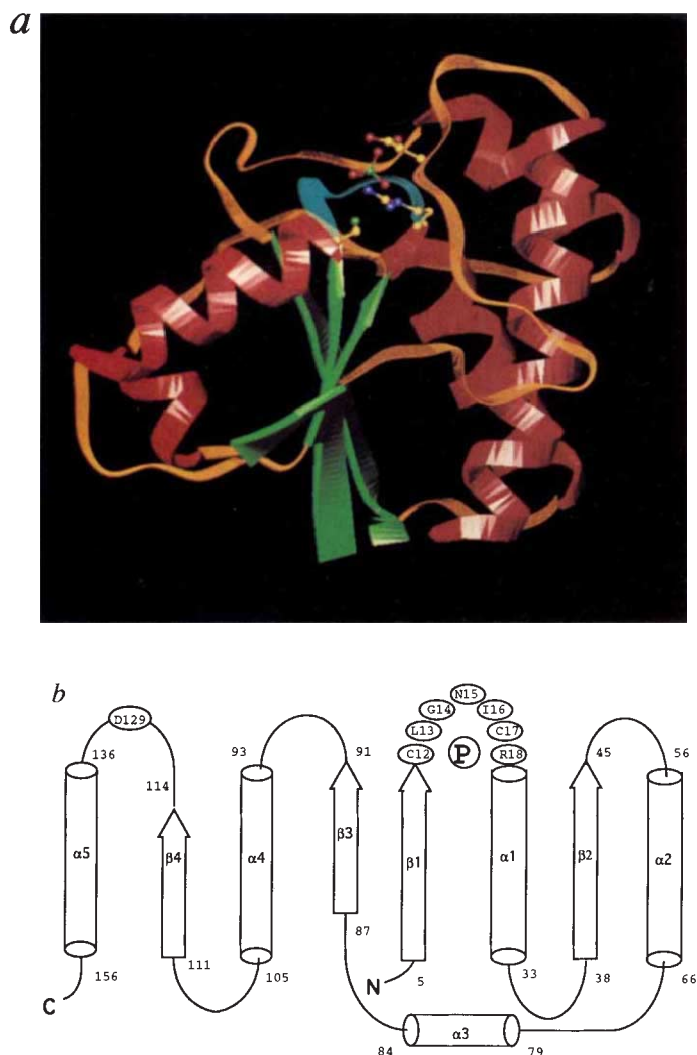


FIG. 1 *a*, The fold of the low-molecular-weight PTPase showing the central twisted β -sheet in green and the surrounding helices in red. A sulphate ion is found in the binding site for the phosphotyrosine substrate at the N-terminal end of helix $\alpha 1$. The loop encircling the sulphate ion is shown in blue. The active-site residues Cys 12, Arg 18 and Asp 129 are also shown. *b*, Schematic diagram of the secondary structure connectivity of the low-molecular-weight PTPase. The phosphate binding site is labelled P.

The catalytic cycle of the protein involves a nucleophilic attack by a cysteine residue, presumably Cys 12, on the phosphorus atom of the substrate, coupled with a release of the tyrosine peptide product^{11,12}. The phospho-cysteinyl residue thereby formed would subsequently be dephosphorylated in a step that is rate-limiting for most substrates¹³. The release of the tyrosine-containing product in the first step of the reaction has been suggested to require an additional electrophile which would enforce the protonation of the leaving tyrosine OH group¹³. We can now consider these suggestions in the light of the structure of the low-molecular-weight PTPase. Cys 12 is localized at the bottom of the active-site cavity, forming a hydrogen bond to a sulphate oxygen (Fig. 3). This cysteine appears to be in an excellent position for a nucleophilic attack on the tyrosine phospho-moiety localized in this pocket. The possible phosphotyrosine-binding positions allowed by the geometry of the active-site crevice strongly support the occurrence of the nucleophilic attack on the opposite side of the leaving tyrosine group, featuring an in-line attack on the phosphotyrosine. This puts Asp 129 in a key position for catalysis at the site of the leaving OH group. Because of its distorted pK_a , Asp 129 could act as a shuttle for the proton required for the leaving tyrosine group. In contrast to many other phosphatases, the low-molecular-weight PTPase does not use metal ions for the activation of the water molecule responsible for the nucleophilic attack of the phospho-cysteinyl intermediate. However, Asp 129 appears to be in a suitable position to serve as a general-base catalyst for the water nucleophile in the dephosphorylation reaction.

The low-molecular-weight PTPase has no notable sequence homology to the large family of homologous PTPases, including Cdc25, human leukocyte common antigen-related protein (LAR) and PTPIB^{14,15}. However, these proteins, as well as the low-molecular-weight PTPase, feature a cysteine residue in the active site that acts as an attacking nucleophile in the phosphotyrosine dephosphorylation, and their reaction mechanisms appear to be related^{12,16}. The phosphate-binding loop of the low-molecular-weight PTPase shows a striking structural similarity to the phosphate-binding loop of the recently described structure

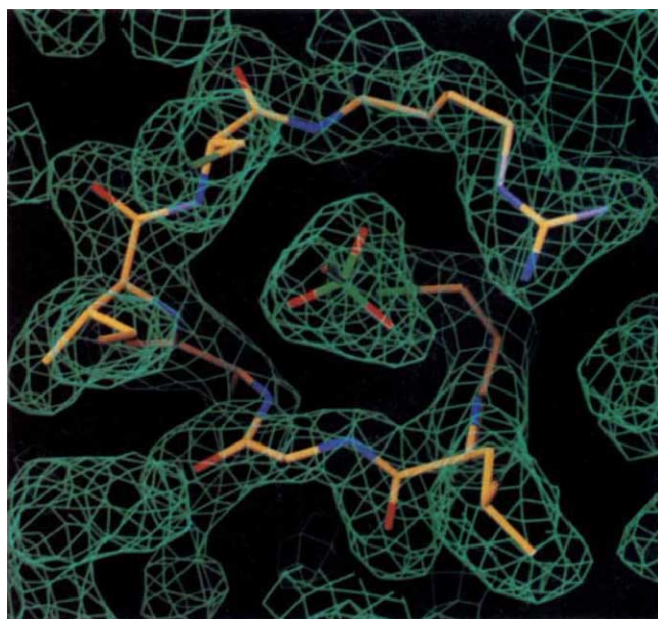


FIG. 2 The electron density map for the active-site region of the low-molecular-weight PTPase. The density clearly shows the bound sulphate ion. The map was calculated by using $(2|F_o| - |F_c|) \exp(iac)$ Fourier coefficients, where F_o and F_c are the observed and calculated structure factors, respectively, and ac is the calculated phase from the final model. The electron density was contoured at 1.1 times the r.m.s. deviation of the map.

TABLE 1 Data collection and refinement statistics

Compound	Resolution (Å)	Unique reflections	Completeness (%)	R_{merge} (%) [*]	Isomorphous difference (%)	Number of sites	Phasing power to 3.0 Å [†]
Native	2.1	9,896	90	6.4			
KAu(CN) ₂	2.5	5,902	90	6.0	19.5	2	1.0
UO ₂ Ac ₂ -1	2.5	4,740	72	5.4	12.0	4	1.3
UO ₂ Ac ₂ -2	2.6	5,327	90	7.9	15.0	4	1.6
UO ₂ Ac ₂ -3	2.5	6,495	97	7.7	20.4	4	1.7
EMTS1 [‡]	2.9	3,620	84	9.8	20.2	3	2.1
EMTS2	3.1	3,253	92	12.1	29.0	3	2.4
EMTS3	2.5	6,213	94	5.6	15.5	3	1.7

Crystals: crystals of the low-molecular-weight PTPase from bovine liver were obtained by using 40% ammonium sulphate, 10 mM DTT in 100 mM acetate buffer, pH 5.5, with a macro-seeding procedure. The lattice is orthorhombic with unit cell dimensions $a=46.3$ Å, $b=62.2$ Å, $c=62.7$ Å, and the space group is $P2_12_12_1$. The asymmetric unit contains one PTPase molecule. Diffraction data: the native and heavy-atom derivative data sets used for the structure determination were collected on a XENTRONICS multiwire area detector and were processed with the program XDS¹⁷. All merging and scaling of the data were done using programs in the CCP4 program suite (SERC, Daresbury, UK). The difference Patterson map for the EMTS1 derivative was easily interpretable, and the major mercury-binding sites were identified and refined with the program MLPHARE (Zbysek Ottonowski, unpublished). The heavy-atom structures for the EMTS, uranyl acetate and AuCN₂ derivatives were completed using difference Fourier techniques. The final mean figure merit was 0.722 to 3.0 Å resolution. The MIR maps were further improved by conventional solvent flattening^{18,19} and SQUASH²⁰. Model building and refinement: skeletonized electron density maps²¹ were used to locate the molecule and to identify the helices and the central β -sheets. The program O was used for all model building²². Initially a polyalanine model of the helices and the central β -sheets was built into the electron density. This first model was subjected to simulated annealing refinement by using X-PLOR²³, and partial model phases were combined with MIR phases. The R -factor from X-PLOR for this model was 42%. Combined maps were used for further iterative model building. After 6 iterations of model building, refinement, phase combination and phase extension, all connectivity errors were located and the sequence was traced. The current model contains 154 (4–157) amino acids, one sulphate ion and 83 water molecules. The current R -value was 15.6%, free R -factor was 20.8% to 2.1 Å resolution and the r.m.s. deviations from ideal bond distances and angles were 0.012 Å and 2.6°, respectively.

^{*} $R_{\text{merge}} = \sum |I_{\text{obs}} - I_{\text{avg}}| / \sum I_{\text{obs}}$ where the summation is over all reflections.

[†] Phasing power = $\text{r.m.s.}(|F_{\text{H}}|_{\text{calc}}) / \text{r.m.s.}(\epsilon)$ where $\text{r.m.s.}(|F_{\text{H}}|_{\text{calc}})$ is the root-mean-square calculated heavy-atom structure factor and $\text{r.m.s.}(\epsilon)$ is the root-mean-square lack of closure error.

[‡] EMTS, Sodium methylmercurithiosalicylate.

of human placenta PTP1B⁹. Although both enzymes are α/β -proteins with a central β -sheet, the ordering of the secondary structure elements is different. The phosphate-binding loops in both proteins include a CXXXXXR motif that forms a cradle-like structure, leaving the arginine and the cysteine residues in close contact with each other. In both proteins, the main-chain amide groups of the loop structures are exclusively exposed into the phosphate-binding site and together with the guanidine

group of Arg 18 they form a compact phosphate-binding motif. However, the interactions stabilizing the loop structure in the two proteins are very different. In addition, the regions of the active site that are possibly involved in the interaction with the peptide substrate differ substantially. In fact they are dominated by basic residues in PTP1B⁹ and by hydrophobic residues in low-molecular-weight PTPase. The active-site crevices in both proteins are fairly deep and, as it has been suggested for PTP1B⁹,

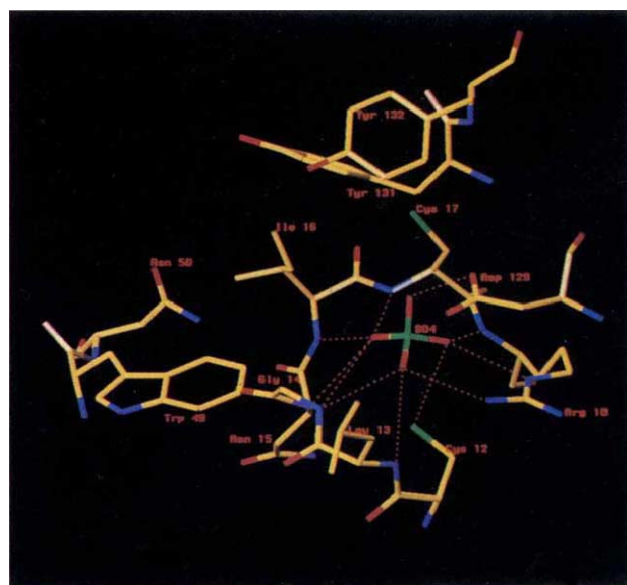
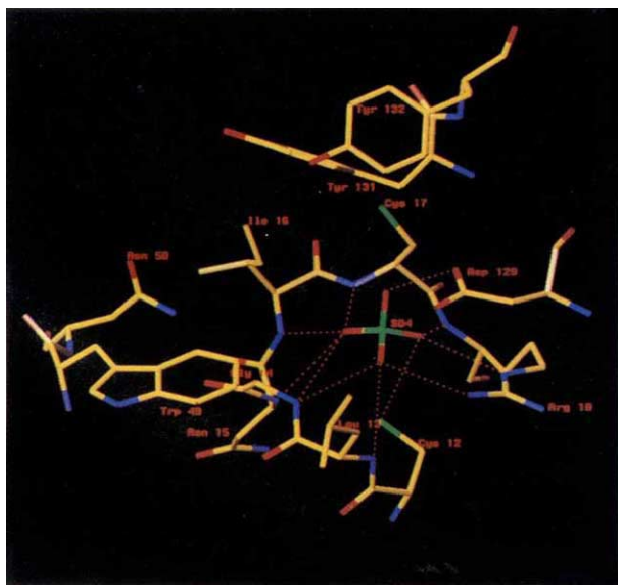


FIG. 3 Stereo diagram of the active site of the low-molecular-weight PTPase. Hydrogen bonds are shown as dashed lines. The sulphate ion forms an extensive hydrogen bonding network to residues 12 to 18, and a hydrogen bond to Asp 129. Cys 12 and Asp 129 are in key positions for catalysis. The distance between the sulphur of the sulphate ion and the sulphur of Cys 12 is 3.8 Å. The hydrogen bond between the

side chain of Asp 129 and the sulphate oxygen is 2.7 Å. Leu 13, Ile 16, Trp 49, Asn 50, Tyr 131, Tyr 132 are residues lining the active-site crevice and may be involved in interactions with the aromatic moiety of the phosphotyrosine-containing substrate, as well as forming a binding site for neighbouring amino acids in the peptide substrate.

this is probably the cause of the inability of the two enzymes to dephosphorylate phosphoserine- and phosphothreonine-containing peptides.

A suitable proton-donating group corresponding to the Asp 129 of the low-molecular-weight PTPase has yet to be identified in PTP1B. Despite the similarities seen in the phosphate-binding sites of the two PTPases, there are sharp differences in the design of their catalytic sites. These differences may be indicative of a history of convergent evolution, highlighting the CXXXXXR loop motif as the common denominator for phosphotyrosine protein phosphatase chemistry. □

Received 13 April; accepted 8 July 1994.

1. Murray, A. W. *Nature* **359**, 599–604 (1992).
2. Camici, G. et al. *J. Biol. Chem.* **264**, 2560–2567 (1989).
3. Chernoff, J. & Li, H. C. *Arch. Biochem. Biophys.* **240**, 135–145 (1985).
4. Ramponi, G. et al. *FEBS Lett.* **250**, 469–473 (1989).
5. Stefani, M. et al. *FEBS Lett.* **325**, 131–134 (1993).
6. Berti, A. *FEBS Lett.* (in the press).
7. Ramponi, G. et al. *Int. J. Cancer* **51**, 652–656 (1992).
8. Ruggiero, M. et al. *FEBS Lett.* **326**, 294–298 (1993).
9. Barford, D., Flint, A. J. & Tonks, N. K. *Science* **263**, 1397–1404 (1994).
10. Su, X.-D. et al. *FEBS Lett.* **343**, 107–108 (1994).
11. Cirri, P. et al. *Eur. J. Biochem.* **214**, 647–657 (1993).
12. Wo, Y. Y. et al. *Biochemistry* **31**, 1712–1721 (1992).
13. Zhang, Z. Y. & Van Etten, R. L. *Biochemistry* **30**, 8954–8959 (1991).
14. Hiraga, A. & Tzuiiki, S. *Adv. Prot. Phosphatases* **6**, 251–264 (1991).
15. Charbonneau, H. & Tonks, N. K. *A. Rev. Cell Biol.* **8**, 463–493 (1992).
16. Guan, K. L. & Dixon, J. E. *J. Biol. Chem.* **266**, 17026–17030 (1991).
17. Kabsch, W. *J. appl. Crystallogr.* **21**, 916–924 (1988).
18. Wang, B. C. *Meth. Enzym.* **115**, 90–112 (1985).
19. Leslie, A. *Acta crystallogr.* **A43**, 134–135 (1987).
20. Zhang, K. Y. J. & Main, P. *Acta crystallogr.* **A46**, 377–381 (1990).
21. Jones, T. A. & Thirup, S. *EMBO J.* **5**, 819–822 (1985).
22. Jones, T. A., Bergdoll, M. & Kjeldgaard, M. in *Crystallographic Computing and Modelling Methods in Molecular Design* (eds Bugg, C. & Ealick, S.) (Springer, New York, 1989).
23. Brünger, T. A., Kuriyan, J. & Karplus, M. *Science* **235**, 458–460 (1987).

ACKNOWLEDGEMENTS. We thank M. Bucciantini for protein purification. This work has been supported by grants from Swedish Natural Science Research Council and the Swedish Cancer Society (P.N.), and from Italian MURST (40%), CNR (Target project on Biotechnology and Bio-instrumentation) and Italian Association for Cancer Research (AIRC). X.-D.S. has been partially supported by the Department of Cell and Molecular Biology, Karolinska Institutet, Stockholm, Sweden.

Premature translational termination triggers mRNA decapping

Denise Muhlrad & Roy Parker

Department of Molecular and Cellular Biology, Life Sciences South, University of Arizona, Tucson, Arizona 85721, USA

THE degradation of messenger RNA in eukaryotic cells is initiated by endonucleolytic cleavage^{1,2} or by shortening of the poly(A) tail^{3–6}, which for some mRNAs activates a deadenylation-dependent decapping reaction⁷. One type of rapid mRNA degradation in eukaryotes is caused by premature termination of translation^{8,9}. This turnover process prevents the translation of aberrant mRNAs^{10,11}, may affect the abundance and splicing pattern of nuclear transcripts^{12,13}, and may be involved in the aetiology of human genetic disease¹⁴. Here we show that premature translational termination in yeast triggers decapping, independent of deadenylation, thereby exposing the transcript to 5'-to-3' degradation. Inactivation of the 5'-to-3' exonuclease reveals an additional 3'-to-5' pathway of mRNA turnover. These observations provide *in vivo* evidence for two new mechanisms of mRNA decay.

To investigate the mechanism of mRNA decay triggered by premature termination of translation, we introduced a nonsense codon at base 103 (codon 22) of the stable *PGK1* mRNA (termed *PGK1_{N103}*). To allow detection of decay products (see later), we also inserted a poly(G) tract into this transcript's 3' untranslated region (UTR) (termed *PGK1_{N103}pG*). Consistent

with earlier results using nonsense mutations¹⁵, the N103 lesion caused a >10-fold increase in the rate of transcript degradation compared with the parental transcript (Fig. 1a, b).

As the degradation of some mRNAs, including *PGK1*, requires shortening of the poly(A) tail^{3,16}, we determined whether the rapid degradation resulting from the N103 nonsense codon required deadenylation. Using the yeast *GAL1* upstream activation sequence (UAS) to induce and repress rapidly the transcription of the *PGK1_{N103}* mRNA, a 'pulse' of transcripts was produced with long poly(A) tails whose decay could be monitored during an ensuing 'chase' (a transcriptional pulse-chase). Two observations indicate that deadenylation is not required for rapid decay of the *PGK1_{N103}* transcript. First, in contrast to the *PGK1* mRNA, which has a temporal lag during which the poly(A) tail is shortened before decay starts (Fig. 2a, e, and ref. 3), the *PGK1_{N103}* transcript begins to decay immediately after transcriptional repression (Fig. 2b, e). Second, insertion of a poly(G) tract (which can form a strong RNA

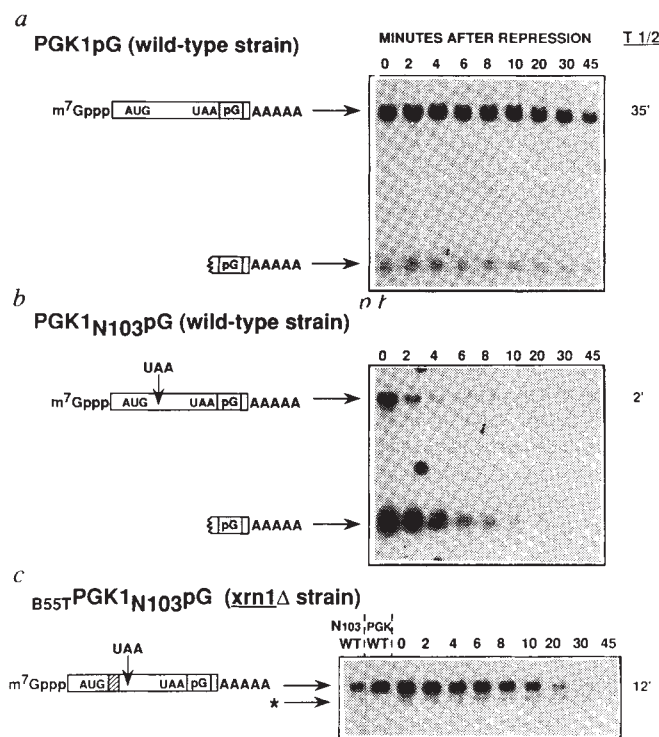


FIG. 1 Decay of *PGK1pG* and *PGK1_{N103}pG* mRNAs after transcriptional repression in cultures grown in galactose. Designations are: pG, 3' UTR poly(G) insertion; boldface UAA, nonsense mutation; striped box, neutral tagging sequence. **a**, **b**, Decay in wild-type yeast strain *yRP582* (*MATa*, *rpb1-1*, *ura3-52*, *leu2*), and **c**, in isogenic *xrn1Δ* strain *yRP689* (*MATa*, *rpb1-1*, *ura3-52*, *leu2*, *xrn1Δ::URA3*). Time points refer to minutes after transcriptional repression. mRNA half-lives are an average of three or more experiments. The smaller band in **a** and **b** represents an mRNA fragment stabilized by the insertion of the poly(G) run that blocks 5'-to-3' decay of the transcripts³. In **c**, the two left lanes are zero-time points from wild-type strains containing *B55T-PGK1_{N103}pG* (labelled N103/WT) or *B55T-PGK1pG* (labelled PGK/WT). The decay intermediate created by 3' to 5' degradation in **c** is indicated by an asterisk.

METHODS. The N103 and B55 mutations were created using the oligonucleotides (5'-GAAGTCAACTTAGATGAAGAC-3' (oRP131)) and (5'-TGGACAGAGATCTTTGAAG-3') respectively. Transcripts were expressed from the *GAL1* UAS present on either plasmid *prP469* (ref. 3) or *prP552*, which contains a *LEU2* gene. The B55 *Bgl*II site and the tagging oligonucleotide inserted in frame at this site (5'-GATCTGTCTCTTTATTCG-ACGACGACGAATT-3') had no effect on mRNA decay (data not shown). Total mRNA (10 μg, prepared as in ref. 22) was run on formaldehyde-agarose gels. Northern blots were probed with an oligonucleotide complementary to the poly(G) sequence and quantitated using a Betascope.

secondary structure capable of blocking yeast exonucleases^{3,17}) into the 3' UTR, termed PGK1_{N103}pG, caused accumulation of a decay product (Fig. 2c, d) that extends from the 5' side of the poly(G) tract to the 3' end of the mRNA (ref. 3, and data not shown). This mRNA fragment was heterogeneous in size, shortened with time, and could be reduced to a uniform length by cleavage with oligo(dT) and RNaseH (Fig. 2d), indicating that this decay product is initially produced with a long poly(A) tail. This is unlike the analogous fragment produced in deadenylation-dependent degradation of the wild-type *PGK1* transcript, which has only an oligo(A) tail (Fig. 2a and ref. 3). We conclude that degradation of the PGK1_{N103}pG transcript must begin before removal of the poly(A) tail.

Because insertion of a poly(G) tract into the 3' UTR causes accumulation of the 3' portion of the PGK1_{N103}pG transcript, degradation of the transcript body must be initiated by nucleolytic cleavage 5' of the 3' UTR to expose the transcript to 5'-to-3' degradation. To determine the site of the initial cleavage, we monitored the decay of PGK1_{N103}pG mRNA in a strain lacking the 5'-to-3' exonuclease encoded by the *XRN1* gene¹⁸, which is required for degradation of mRNAs that are decapped after deadenylation⁷. If this 5'-3' exonuclease also functions in the degradation of the PGK1_{N103}pG mRNA, then an *xrn1*Δ strain should accumulate the products of the initial nucleolytic cleavage, the substrates for Xrn1p. In the *xrn1*Δ strain, the full-length transcript with the N103 lesion is stabilized 6-fold (Fig. 1c, upper

band), suggesting that the nucleolytic event triggered by the non-sense codon must be near the 5' end of the transcript.

To determine the site of cleavage, we examined the decay of the PGK1_{N103}pG transcript in the *xrn1*Δ strain by a transcriptional pulse-chase. To allow the 5' end of the PGK1_{N103}pG transcript to be distinguished from endogenous *PGK1* transcripts, we inserted an oligonucleotide near the 5' end (referred to as B55T, Fig. 2f) that has no effect on mRNA decay (data not shown). RNA from zero and 10-min time points was separated into capped and uncapped fractions using antisera directed against the cap structure and then analysed by northern blotting. In wild-type strains, only capped transcripts are detected (Fig. 3a, lanes 2 and 5); in the *xrn1*Δ strain, some B55T-PGK1_{N103}pG transcripts are uncapped (lanes 9 and 12), and by the 10-min time point these predominate (lane 12). As B55T-PGK1_{N103}pG transcripts that accumulated in the *xrn1*Δ strain lacked only the extreme 5' end of the mRNA, we conclude that premature translational termination triggers decapping, independently of deadenylation, by cleavage at, or near, the 5' end of the transcript (Fig. 4).

Based on primer extension assays (Fig. 3b), the uncapped nonsense mRNAs present at 10 and 30 min in the *xrn1*Δ strain are mostly shortened at their 5' ends by two nucleotides relative to the major transcriptional start site (Fig. 3b); on longer exposure, some transcripts are detected that are shortened by one nucleotide at the 5' end (data not shown). This suggests that the

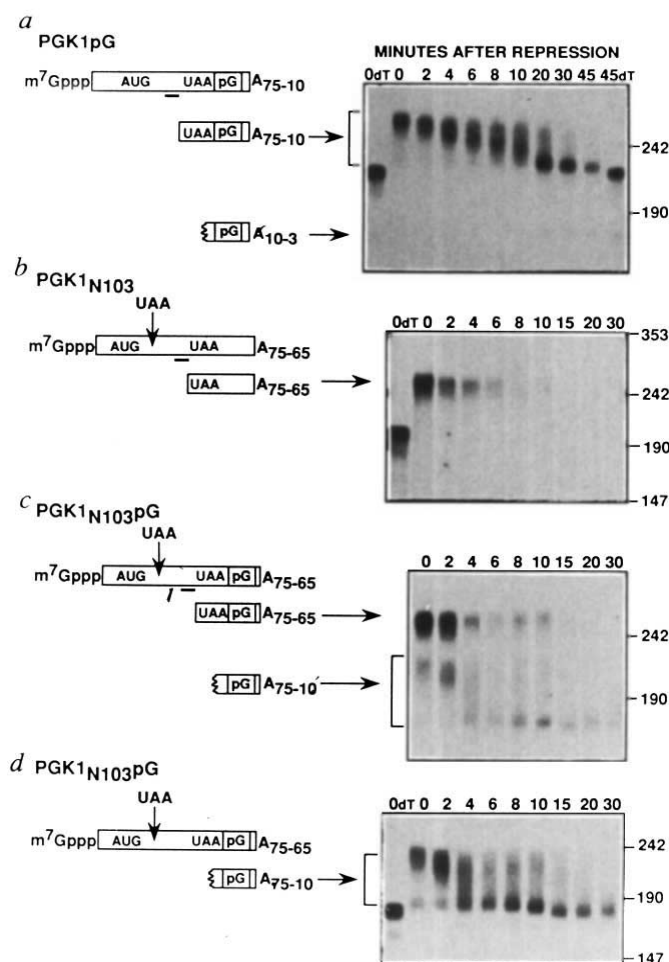
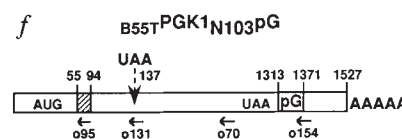
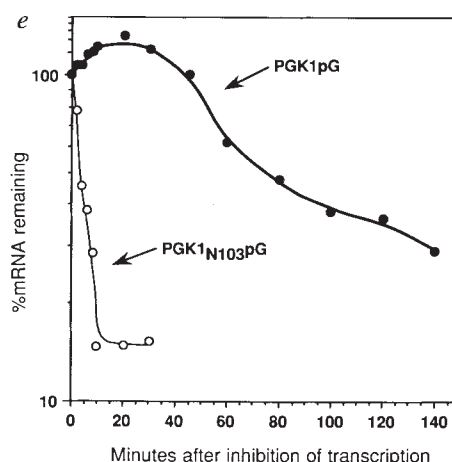


FIG. 2 Analysis of deadenylation and decay of the PGK1_{N103}pG mRNA. a-d, Transcriptional pulse-chase analyses of the relationship between deadenylation and decay in strain yRP582, where the numbers after the 'A' signify poly(A) tail length. In a-c, mRNAs were cleaved with RNaseH and



oligo oRP70 (5'-CGGATAAGAAAGCAACACCTGG-3') (underlined), to allow resolution of poly(A) tail sizes. In d, no RNaseH treatment was used. Lanes labelled 'dT' contain an mRNA sample that was also cleaved with RNaseH and oligo(dT) to remove the poly(A) tail. DNA sizes (bp) are indicated on the right. Blots were probed as for Fig. 1 and the bands quantified on a Betascope. The construct in b contains a small insertion of poly(G) in its 3' UTR which is insufficient to form a stable secondary structure but allows unique hybridization to the probe. e, Average half-life of PGK1pG and PGK1_{N103}pG mRNAs from several experiments. f, B55T-PGK1_{N103}pG mRNA: numbers above indicate base position from 5' end; striped box, neutral tag sequence; pG, 3' UTR poly(G) insertion; boldface UAA, nonsense mutation. Oligonucleotides (o) are indicated by arrows; their sequences are given in the figure legends.

METHODS. Transcriptional pulse-chases, RNaseH cleavages, and northern blotting have all been described^{3,16}.

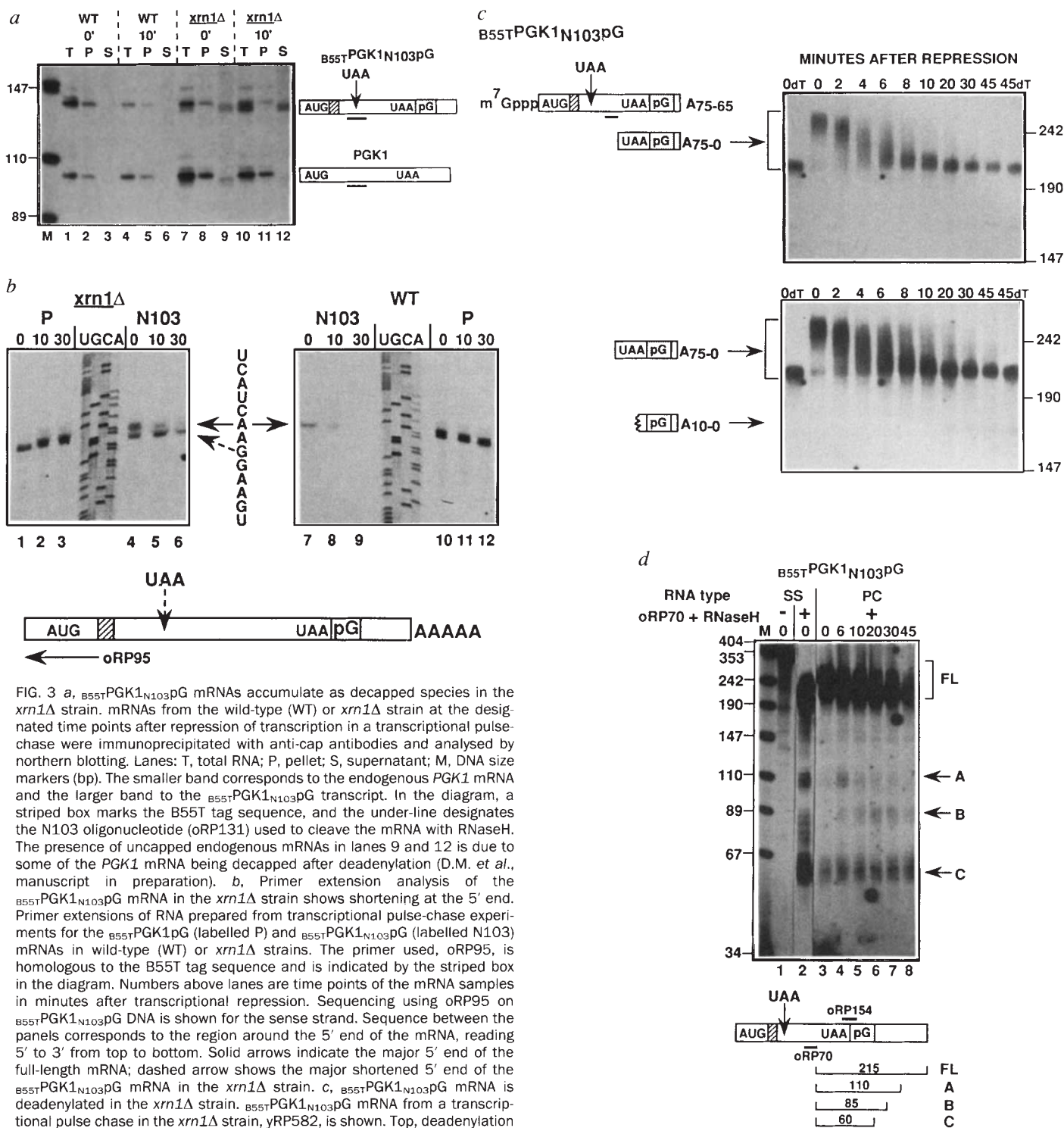


FIG. 3 *a*, B55T-PGK1N103pG mRNAs accumulate as decapped species in the *xrn1Δ* strain. mRNAs from the wild-type (WT) or *xrn1Δ* strain at the designated time points after repression of transcription in a transcriptional pulse-chase were immunoprecipitated with anti-cap antibodies and analysed by northern blotting. Lanes: T, total RNA; P, pellet; S, supernatant; M, DNA size markers (bp). The smaller band corresponds to the endogenous PGK1 mRNA and the larger band to the B55T-PGK1N103pG transcript. In the diagram, a striped box marks the B55T tag sequence, and the under-line designates the N103 oligonucleotide (oRP131) used to cleave the mRNA with RNaseH. The presence of uncapped endogenous mRNAs in lanes 9 and 12 is due to some of the PGK1 mRNA being decapped after deadenylation (D.M. *et al.*, manuscript in preparation). *b*, Primer extension analysis of the B55T-PGK1N103pG mRNA in the *xrn1Δ* strain shows shortening at the 5' end. Primer extensions of RNA prepared from transcriptional pulse-chase experiments for the B55T-PGK1N103pG (labelled P) and B55T-PGK1N103pG (labelled N103) mRNAs in wild-type (WT) or *xrn1Δ* strains. The primer used, oRP95, is homologous to the B55T tag sequence and is indicated by the striped box in the diagram. Numbers above lanes are time points of the mRNA samples in minutes after transcriptional repression. Sequencing using oRP95 on B55T-PGK1N103pG DNA is shown for the sense strand. Sequence between the panels corresponds to the region around the 5' end of the mRNA, reading 5' to 3' from top to bottom. Solid arrows indicate the major 5' end of the full-length mRNA; dashed arrow shows the major shortened 5' end of the B55T-PGK1N103pG mRNA in the *xrn1Δ* strain. *c*, B55T-PGK1N103pG mRNA is deadenylated in the *xrn1Δ* strain. B55T-PGK1N103pG mRNA from a transcriptional pulse chase in the *xrn1Δ* strain, yRP582, is shown. Top, deadenylation of mRNA; bottom, overexposure of the same northern blot, allowing visualization of the decay fragment. To the left are shown mRNA species with the poly(A) tail length for each band. Boldface UAA denotes the N103 mutation, the underline the oligonucleotide oRP70 (Fig. 2) used to cleave the mRNA with RNaseH, and dT samples for which oligo(dT) was also used to remove poly(A) tails. DNA size markers are given on the right (bp). The northern blot was probed and quantified as for Fig. 1. *d*, Nonsense mRNAs accumulate 3'-to-5' exonucleolytic decay products in the *xrn1Δ* strain. RNA was prepared from an *xrn1Δ* strain expressing the B55T-PGK1N103pG mRNA either from a steady-state culture (SS, lanes 1 and 2) or from a transcriptional pulse-chase experiment (PC, lanes 3–8). The mRNA was cleaved with oligonucleotide oRP70 (Fig. 2) and RNaseH to resolve small size differences at the 3' end, and analysed by northern blotting. Lane M, DNA size markers (bp). Time points in minutes after transcriptional repression during the pulse-chase are shown at the top. Total mRNA loaded was 10 µg in the first two lanes and 25 µg in all other lanes. The bracket labelled FL indicates the mRNA's full-length 3' end and bands A, B and C (arrowed) are transcripts

shortened at their 3' end. B and C represents molecules shortened to the 3' side of the poly(G) tract, present even at zero time because of some preinduction of the transcript. Underneath are shown the mRNA, oligonucleotide oRP70 used to cleave the mRNA, and oRP154 (5'-CCCCCAATTCCTTCGATTTC-3') used to probe the northern specifically for the 5' side of the poly(G) tract. Fragment sizes (nucleotides) are approximate. **METHODS.** *a*, Total RNA (10 µg) from each time point was cleaved with RNaseH and oligonucleotide oRP131 (Fig. 1). Immunoprecipitations were done as described^{7,23}. 6% Polyacrylamide/7M urea gels were electrophoresed, blotted and probed with a riboprobe synthesized from a pGEM-4Z plasmid (Promega) containing B55T-PGK1N103 from the 5' end of the gene through to the KpnI site within the PGK1 coding region, then linearized with EcoRI and transcribed with T7 RNA polymerase. *b*, The primer oRP95 (5'-GATCAATTCGTCGTCGAATAAGAAAGACAA-3') was kinased and used to extend 25 µg total mRNA from each time point by standard methods.

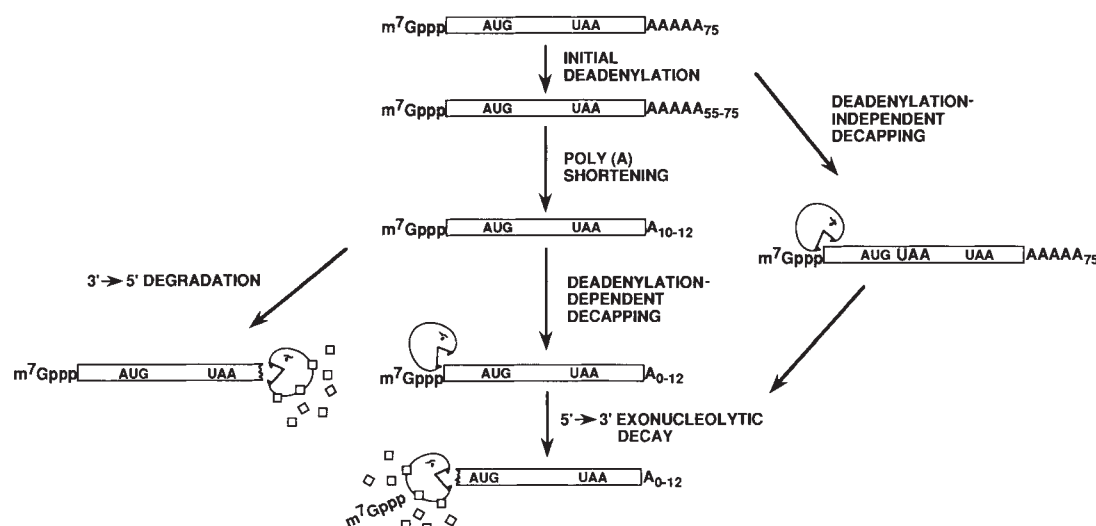


FIG. 4 Relationships of different pathways of mRNA degradation. The two new pathways of mRNA decay described in the text are shown, together with their mechanistic relationship to a general pathway of

mRNA decay in which deadenylation triggers mRNA decapping⁷. Not included are mechanisms of mRNA degradation initiated by endonucleolytic cleavage.

site of decapping could lie within a phosphodiester bond near the 5' end, although one or two nucleotides may be removed from the 5' end following decapping. This is in contrast to *MFA2* mRNA, which when decapped remains full-length⁷.

Analysis of the 3' ends of the *B55T*PGK1_{N103}pG transcripts in the *xrn1Δ* strain indicates that there are one or more additional 5'-to-3' exonucleases that can substitute, albeit inefficiently, for Xrn1p. The critical observation is that in the *xrn1Δ* strain some of the fragment trimmed to the 5' side of the poly(G) insertions is detected after long exposure (Fig. 3c).

Analysis of the *B55T*PGK1_{N103}pG transcripts in the *xrn1Δ* strain also reveals an additional pathway of mRNA decay in which the transcript is degraded 3' to 5' following deadenylation. An mRNA fragment of ~1.3 kilobases (kb) accumulates, which is slightly smaller than the 1.45-kb full-length transcript (Fig. 1c, lower band). Examination of the 3' end of these transcripts by RNase protection (data not shown) and northern blotting identified a set of new 3' ends, the major end being located at the 3' side of the poly(G) tract (Fig. 3d, bands A, B and C in lane 2). Moreover, in a transcriptional pulse-chase experiment small amounts of these mRNA fragments were detected after deadenylation (Fig. 3d, bands A, B and C in lanes 3–8). We

conclude that in the absence of Xrn1p, this mRNA can be degraded from the 3' end after deadenylation, exo- and/or endonucleolytically. This 3'-5' decay occurs slowly on *B55T*PGK1_{N103}pG transcripts and is therefore observed only when the principal 5'-3' exonuclease is inactivated. But the existence of a 3'-5' mechanism of degradation indicates that this may be the primary mode of turnover for some transcripts.

The rapid degradation induced by premature termination of translation is unique in that decapping of the transcript is the first nucleolytic step to occur. In contrast, decapping of the *MFA2* mRNA requires prior deadenylation⁷, suggesting that an interaction between the 5' cap and the 3' poly(A) tail prevents decapping. Interactions between termini have also been proposed on the basis of effects of the poly(A) tail and 3' UTR sequences^{19,20} on translation initiation and of the presence of circular polysomes in electron micrographs²¹. Premature termination of translation might activate deadenylation-independent decapping by disrupting or preventing the formation of this 5'-3' interaction. Moreover, these results imply that in addition to early nonsense codons, some sequences which dictate rapid mRNA decay will activate deadenylation-independent decapping. □

Received 13 April; accepted 8 July 1994.

1. Brown, B. D., Zipkin, I. D. & Harland, R. M. *Genes Dev.* **7**, 1620–1631 (1993).
2. Stoeckle, M. *Nucleic Acids Res.* **20**, 1123–1127 (1992).
3. Decker, C. J. & Parker, R. *Genes Dev.* **7**, 1632–1643 (1993).
4. Brewer, G. & Ross, J. *Molec. cell. Biol.* **8**, 1697–1708 (1988).
5. Shyu, A. B., Belasco, J. G. & Greenberg, M. E. *Genes Dev.* **5**, 221–234 (1991).
6. Wilson, T. & Triesman, R. *Nature* **336**, 396–399 (1988).
7. Muhrad, D., Decker, C. J. & Parker, R. *Genes Dev.* **8**, 855–866 (1994).
8. Losson, R. & Lacroute, F. *Proc. natn. Acad. Sci.* **76**, 5134–5137 (1979).
9. Maquet, L. E., Kinniburgh, A. J., Rachmilewitz, E. A. & Ross, J. *Cell* **27**, 543–553 (1981).
10. He, F., Peltz, S. W., Donahue, J. L., Rosbash, M. & Jacobson, A. *Proc. natn. Acad. Sci.* **90**, 7034–7038 (1993).
11. Pulak, R. & Anderson, P. *Genes Dev.* **7**, 1885–1897 (1993).
12. Cheng, J. & Maquet, L. E. *Molec. cell. Biol.* **13**, 1892–1902 (1993).

13. Naeger, L. K., Schoborg, R. V., Zhao, Q., Tullis, G. E. & Pintel, D. J. *Genes Dev.* **6**, 1107–1119 (1992).
14. Dietz, H. C. et al. *Science* **259**, 680–682 (1993).
15. Peltz, S. W., Brown, A. H. & Jacobson, A. *Genes Dev.* **7**, 1737–1754 (1993).
16. Muhrad, D. & Parker, R. *Genes Dev.* **6**, 2100–2111 (1992).
17. Vreken, P. & Raue, H. A. *Molec. cell. Biol.* **12**, 2986–2996 (1992).
18. Stevens, A. & Maupin, M. K. *Arch. Biochem. Biophys.* **252**, 339–347 (1987).
19. Munroe, D. & Jacobson, A. *Gene* **91**, 151–158 (1990).
20. Jackson, R. J. & Standart, N. *Cell* **62**, 15–24 (1990).
21. Christensen, A. K., Kahn, L. E. & Bourne, C. M. *Am. J. Anat.* **178**, 1–10 (1987).
22. Caponigro, G., Muhrad, D. & Parker, R. *Molec. cell. Biol.* **13**, 5141–5148 (1993).
23. Yang, H., Moss, M. L., Lund, E. & Dahlberg, J. E. *Molec. cell. Biol.* **12**, 1553–1560 (1992).

ACKNOWLEDGEMENTS. We thank C. Dieckmann, J. Little, and members of our laboratory for their comments on the manuscript. This work was supported by an NIH grant to R.P.

PLEASE follow these guidelines so that your manuscript may be handled expeditiously.

Nature is an international journal covering all the sciences. Contributors should therefore bear in mind those readers who work in other fields and those for whom English is a second language, and write clearly and simply, avoiding unnecessary technical terminology. Space in the journal is limited, making competition for publication severe. Brevity is highly valued. One printed page of *Nature*, without display items, contains about 1,300 words.

Manuscripts are selected for publication according to editorial assessment of their suitability and reports from independent referees. They can be sent to London or Washington and should be addressed to the Editor. Manuscripts may be dealt with in either office, depending on the subject matter, and will where necessary be sent between offices by overnight courier. High priority cannot be given to pre-submission enquiries; in urgent cases they can be made in the form of a one-page fax. All manuscripts are acknowledged on receipt but fewer than half are sent for review. Those that are not reviewed are returned as rapidly as possible so that they may be submitted elsewhere without delay. Contributors may suggest reviewers; limited requests for the exclusion of specific reviewers are usually heeded. Manuscripts are usually sent to two or three reviewers, chosen for their expertise rather than their geographical location. Manuscripts accepted for publication are processed from the London office.

Nature requests authors to deposit sequence and crystallographic data in the databases that exist for this purpose.

Once a manuscript is accepted for publication, contributors will receive proofs in about 4 weeks. *Nature's* staff will edit manuscripts with a view to brevity and clarity, so contributors should check proofs carefully. Manuscripts are generally published 2–3 weeks after receipt of corrected proofs. *Nature* does not exact page charges. Contributors receive a reprint order form with their proofs; reprint orders are processed after the manuscript is published and payment received.

Categories of paper

Review Articles survey recent developments in a field. Most are commissioned but suggestions are welcome in the form of a one-page synopsis addressed to the Reviews Coordinator. Length is negotiable in advance.

Progress articles review particularly topical developments for a nonspecialist readership. They do not exceed 4 pages in length. Suggestions may be made to the Reviews Coordinator in the form of a brief synopsis.

Articles are research reports whose conclusions are of general interest and which are sufficiently rounded to be a substantial advance in understanding. They should not have more than 3,000 words of text (not including figure legends) or more than six display items and should not occupy more than five pages of *Nature*.

Articles start with a heading of 50–80 words written to advertise their content in general terms, to which editors will pay particular attention. The heading does not usually contain numbers, abbreviations or measurements. The introduction to the study is contained in the first two or three paragraphs of the article, which also briefly summarize its results and implications. Articles have fewer than 50 references and may contain a few short subheadings.

Letters to *Nature* are short reports of outstanding novel findings whose implications are general and important enough to be of interest to those outside the field. Letters should have 1,000 or fewer words of text and four or fewer display items. The first paragraph describes, in not more than 150 words and without the use of abbreviations, the background, rationale and chief conclusions of the study for the particular benefit of non-specialist readers. Letters do not have subheadings and contain fewer than 30 references.

Commentary articles deal with issues in, or arising from, research that are also of interest to readers outside research.

News and Views articles inform nonspecialist readers about new scientific advances, sometimes in the form of a conference report. Most are commissioned but proposals can be made in advance to the News and Views Editor.

Scientific Correspondence is for discussion of topical scientific matters, including those published in *Nature*, and for miscellaneous contributions. Priority is given to letters of fewer than 500 words.

Preparation of manuscripts

All manuscripts should be typed, double-spaced, on one side of the paper only. An original and four copies are required, each accompanied by artwork. If photographs are included, five sets of originals are required; for line drawings, one set of originals and four good-quality photocopies are acceptable. Reference lists, figure legends and tables should all be on separate sheets, all of which should be double-spaced and numbered. Three copies of relevant manuscripts in press or submitted for publication elsewhere should be included with submitted manuscripts, clearly marked as such. Five copies of revised and resubmitted manuscripts, labelled with their manuscript numbers, are required, together with five copies of a letter detailing the changes made.

Titles are brief and simple. Active verbs, numerical values, abbreviations and punctuation are to be avoided. Titles should contain one or two key words for indexing purposes.

Artwork should be marked individually and clearly with the author's name and, when known, the manuscript number. Ideally, no figure should be larger than 28 by 22 cm. Figures with several parts are to be avoided and are permitted only if the parts are closely related, either experimentally or logically. Unlettered originals of photographs should be provided. Suggestions for cover illustrations, with captions and labelled with the manuscript number, are welcome. Original artwork is returned when a manuscript cannot be published.

Protein/nucleotide sequences should ideally be in the three-letter and not the single-letter code for amino acids. One column width of *Nature* can accommodate 20 amino acids or 60 base pairs.

Colour artwork. A charge of £500 per page is made as a contribution towards the cost of reproducing colour figures. Inability to pay these costs will not prevent publication of essential colour figures if the circumstances are explained. Proofs of colour artwork may be sent to contributors under separate cover from their galley proofs.

Figure legends should not exceed 300 words and ideally should be shorter. The figure is described first, then, briefly, the method. Reference to a method published elsewhere is preferable to a full description. Methods are not described in the text.

References are numbered sequentially as they appear in the text, followed by those in tables and finally by those in figure legends. Only papers published or in the press are numbered and included in the reference list. All other forms of reference should be cited in the text as a personal communication, manuscript submitted or in preparation. Text is not included in reference lists. References are abbreviated according to the *World List of Scientific Periodicals* (Butterworths, London, 1963–65). The first and last page numbers are included; reference to books includes publisher, place and date.

Abbreviations, symbols, units and Greek letters should be identified the first time they are used. Acronyms should be avoided whenever possible and, if used, defined. Footnotes are not used in the text.

Acknowledgements are brief and appear after the reference list; grant and contribution numbers are not allowed.

Supplementary information is material relevant to Articles or Letters which cannot, for lack of space, be published in full, but which is available from *Nature* on request.

Submission. Manuscripts can be sent to the Editor at 4 Little Essex Street, London WC2R 3LF, UK or at 1234 National Press Building, Washington, DC 20045, USA. A telephone and fax number should be included. Manuscripts or proofs sent by air courier to London should be declared as 'manuscripts' and 'value \$5' to prevent the imposition of import duty and value-added tax.